

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 12:36:32 ; Search time 46 Seconds
(without alignments)
3114.027 Million cell updates/sec

Title: US-09-455-486-6
Perfect score: 2351
Sequence: 1 MESISMMSPKSLSETCLPN.....ALVPSIVILLQLCRYPD 454

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2290	97.4	490	4 Q81UE7	Q81ue7 homo sapien
2	2284	97.2	490	4 Q8NFT2	Q8nft2 homo sapien
3	2239.5	95.3	488	11 Q8BMB6	Q8bmb6 mus musculu
4	1318	56.1	488	11 Q8C5F0	Q8c5f0 mus musculu
5	1316	56.0	488	11 Q99P41	Q99p41 rattus norv
6	1316	56.0	488	11 Q8C159	Q8c159 mus musculu
7	1316	56.0	526	11 Q80ZP3	Q80zf3 mus musculu
8	1306	55.6	514	11 Q924Z1	Q924z1 mus musculu
9	1272	54.1	488	4 Q86SF6	Q86sf6 homo sapien
10	1272	54.1	498	4 Q7Z389	Q7z389 homo sapien
11	1269	54.0	488	4 Q9NVB5	Q9nvb5 homo sapien
12	1252.5	53.3	487	4 Q8NEW6	Q8new6 homo sapien
13	1245.5	53.0	487	4 Q8TF03	Q8tf03 homo sapien
14	1074.5	45.7	456	4 Q8TDP3	Q8tdp3 homo sapien
15	1064.5	45.3	470	11 Q923B6	Q923b6 mus musculu
16	1057.5	45.0	470	11 Q91W31	Q91w31 mus musculu

17	1041.5	44.3	474	11 Q91ZE8	Q91ze8 mus musculu
18	956.5	40.7	464	11 Q7TP88	Q7tp88 rattus norv
19	728	31.0	338	6 Q9GL50	Q9gl50 sus scrofa
20	715	30.4	339	11 Q924Z2	Q924z2 mus musculu
21	707	30.1	339	11 Q9CWR7	Q9cwr7 mus musculu
22	704	29.9	339	11 Q924J9	Q924j9 mus musculu
23	628	26.7	283	4 Q9H5R1	Q9h5r1 homo sapien
24	541	23.0	283	4 Q8WMB0	Q8wmb0 homo sapien
25	324	13.8	143	4 Q9H7Y1	Q9h7y1 homo sapien
26	171.5	7.3	208	16 Q8NLU6	Q8nlue corynebacte
27	170	7.2	208	16 Q8XQ52	Q8xqs2 ralsstonia s
28	164.5	7.0	234	16 Q82AX0	Q82ax0 streptomyce
29	160	6.8	198	16 Q93OK7	Q93ok7 rhizobium m
30	159.5	6.8	211	16 Q8YK44	Q8yk44 anabaena sp
31	156.5	6.7	239	16 Q9RKR6	Q9rkr6 streptomyce
32	154.5	6.6	225	2 Q8GFG4	Q8gfg4 rhodococcus
33	151	6.4	200	16 Q98C62	Q98c62 rhizobium 1
34	149.5	6.4	218	2 Q8GFG3	Q8gfg3 rhodococcus
35	149.5	6.4	221	16 Q82Q14	Q82q14 streptomyce
36	147.5	6.3	226	2 Q8GKH6	Q8gkh6 rhodococcus
37	147	6.3	213	17 Q29059	Q29059 archaeoglob
38	144.5	6.1	226	2 Q9AH05	Q9ah05 rhodococcus
39	143.5	6.1	223	17 Q8TYQ9	Q8tyq9 methanopyru
40	142	6.0	212	17 Q29370	Q29370 archaeoglob
41	142	6.0	224	1 Q59661	Q59661 methanobact
42	137.5	5.8	199	2 Q9F418	Q9f418 mycobacteri
43	136.5	5.8	217	2 Q83VC1	Q83vc1 streptomyce
44	136.5	5.8	217	16 Q9JN78	Q9jnt8 streptomyce
45	135.5	5.8	203	2 Q8GFG1	Q8gfg1 rhodococcus

ALIGNMENTS

RESULT 1	ID	Q81UE7	PRELIMINARY;	PRT,	490 AA.
AC	Q81UE7	Q81UE7	Q81UE7	Q81UE7	Q81UE7
DT	01-MAR-2003	(TREMBlrel. 23, Created)			
DT	01-MAR-2003	(TREMBlrel. 23, Last sequence update)			
DT	01-JUN-2003	(TREMBlrel. 24, Last annotation update)			
DE	STAMP1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_Taxid=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Prostate;				
RX	MEDLINE=22229309; PubMed=12095985;				
RA	Korkmaz K.S., Elbi C.C., Korkmaz C.G., Loda M., Hager G.L.,				
RA	Saatcioglu F.;				
RT	"Molecular cloning and characterization of STAMP1, a highly prostate				
RT	specific six-trans-membrane protein that is overexpressed in prostate				
RT	cancer.";				
RL	J. Biol. Chem. 277:36689-36696(2002).				
DR	EMBL; AY008445; AAG32149.1;				
DR	GO; GO:0006118; P:electron transport; IEA.				
DR	InterPro; IPR004455; NADPoxred_F420.				
DR	Pfam; PF03807; F420_oxidored; 1.				
SO	SEQUENCE 490 AA; 56027 MW; D1971A84880F7E51 CRC64;				

Query Match 97.4%; Score 2290; DB 4; Length 490;
Best Local Similarity 99.8%; Pred. No. 4.1e-176;
Matches 443; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	MESISMMSPKSLSETCLPNNGINGIKDARKVTGVIGSGDFAKSLTIRLRGYNVIGS	60
DB	1	MESISMMSPKSLSETCLPNNGINGIKDARKVTGVIGSGDFAKSLTIRLRGYNVIGS	60
QY	61	RNPKFASFPFHVVDVTHHEDALTKNIIFVAIHREHYTSLMDLRHLLVGKILIDVSNM	120
DB	61	RNPKFASFPFHVVDVTHHEDALTKNIIFVAIHREHYTSLMDLRHLLVGKILIDVSNM	120

QY 121 RINQYPESNAEYLAIFPDSLIVGFNVSAWALQIGPKDASRQVYICSNNIQARQVIE 180
Db 121 RINQYPESNAEYLAIFPDSLIVGFNVSAWALQIGPKDASRQVYICSNNIQARQVIE 180
QY 181 LARQNFIPIDLGSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
Db 181 LARQNFIPIDLGSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
QY 241 RNOQSDFFYKIPIEIVNKTLPVATLLSLVYLAGLAAAYQLYGTYRRFPFWLETWLQ 300
Db 241 RNOQSDFFYKIPIEIVNKTLPVATLLSLVYLAGLAAAYQLYGTYRRFPFWLETWLQ 300
QY 301 CRKQGLISFFPAMVHVAYSCLPMRSEERYFLNMAYQOVHANIENSNWNEEVRRIEMY 360
Db 301 CRKQGLISFFPAMVHVAYSCLPMRSEERYFLNMAYQOVHANIENSNWNEEVRRIEMY 360
QY 361 ISFGIMSLGLSLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
Db 361 ISFGIMSLGLSLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
QY 421 EBYRYFTPPNPFVLAIVLPISIVIL 444
Db 421 EBYRYFTPPNPFVLAIVLPISIVIL 444

RESULT 2

Q8NFT2 PRELIMINARY; PRT; 490 AA.
AC Q8NFT2; 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Six-transmembrane epithelial antigen of prostate 2.
GN STEAP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Poraka K.P., Helenius M.A., Visakorpi T.;
RT "Cloning and characterization of a novel six-transmembrane protein
STEAP2, down-regulated in androgen-independent prostate cancer.";
RT Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF455138; AAN04080.1; -.
DR Genew; HGNC:17885; STEAP2.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR004455; NADPoxred F420.
DR Pfam; PF03807; F420_oxidored; I.
KW Transmembrane.
SQ SEQUENCE 490 AA; 55961 MW; F2E9C30CDACCCEFB1 CRC64;

Query Match 97.2%; Score 2284; DB 4; Length 490;
Best Local Similarity 99.8%; Pred. No. 1.2e-175;
Matches 443; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MESISMGSPPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLIRCGYHVVIGS 60
Db 1 MESISMGSPPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLIRCGYHVVIGS 60
QY 61 RNPKFASEFFPHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVSKILIDVSNM 120
Db 61 RNPKFASEFFPHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVSKILIDVSNM 120
QY 121 RINQYPESNAEYLAIFPDSLIVGFNVSAWALQIGPKDASRQVYICSNNIQARQVIE 180
Db 121 RINQYPESNAEYLAIFPDSLIVGFNVSAWALQIGPKDASRQVYICSNNIQARQVIE 180
QY 181 LARQNFIPIDLGSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
Db 181 LARQNFIPIDLGSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240

Db 181 LARQNFIPIDLGSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
QY 241 RNOQSDFFYKIPIEIVNKTLPVATLLSLVYLAGLAAAYQLYGTYRRFPFWLETWLQ 300
Db 241 RNOQSDFFYKIPIEIVNKTLPVATLLSLVYLAGLAAAYQLYGTYRRFPFWLETWLQ 300
QY 301 CRKQGLISFFPAMVHVAYSCLPMRSEERYFLNMAYQOVHANIENSNWNEEVRRIEMY 360
Db 301 CRKQGLISFFPAMVHVAYSCLPMRSEERYFLNMAYQOVHANIENSNWNEEVRRIEMY 360
QY 361 ISFGIMSLGLSLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
Db 361 ISFGIMSLGLSLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
QY 421 EBYRYFTPPNPFVLAIVLPISIVIL 444
Db 421 EBYRYFTPPNPFVLAIVLPISIVIL 444

RESULT 3

Q8BWB6 PRELIMINARY; PRT; 489 AA.
AC Q8BWB6; 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE weakly similar to tumor suppressor PHYE.
GN 4921538B17RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RT Nature 420:563-573(2002).
RL EMBL; AK052981; BAC35230.1; -.
DR MGD; MGI:1921301; 4921538B17RIK.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR004455; NADPoxred F420.
DR Pfam; PF03807; F420_oxidored; I.
SQ SEQUENCE 489 AA; 55760 MW; 98CD63D59DDDF24C CRC64;

Query Match 95.3%; Score 2239.5; DB 11; Length 489;
Best Local Similarity 97.3%; Pred. No. 4.8e-172;
Matches 432; Conservative 9; Mismatches 2; Indels 1; Gaps 1;

QY 1 MESISMGSPPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLIRCGYHVVIGS 60
Db 1 MESISMGSPPKSL-ETFLPNGINGIKDARKVTGVIGSGDFAKSLTIRLIRCGYHVVIGS 59
QY 61 RNPKFASEFFPHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVSKILIDVSNM 120
Db 60 RNPKFASEFFPHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVSKILIDVSNM 119
QY 121 RINQYPESNAEYLAIFPDSLIVGFNVSAWALQIGPKDASRQVYICSNNIQARQVIE 180
Db 120 RINQYPESNAEYLAIFPDSLIVGFNVSAWALQIGPKDASRQVYICSNNIQARQVIE 179
QY 181 LARQNFIPIDLGSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
Db 180 LARQNFIPIDLGSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 239
QY 241 RNOQSDFFYKIPIEIVNKTLPVATLLSLVYLAGLAAAYQLYGTYRRFPFWLETWLQ 300
Db 240 RNOQSDFFYKIPIEIVNKTLPVATLLSLVYLAGLAAAYQLYGTYRRFPFWLETWLQ 299
QY 301 CRKQGLISFFPAMVHVAYSCLPMRSEERYFLNMAYQOVHANIENSNWNEEVRRIEMY 360
Db 301 CRKQGLISFFPAMVHVAYSCLPMRSEERYFLNMAYQOVHANIENSNWNEEVRRIEMY 360

DB 300 CRKQLGLSFFFAVHVAYSLCLPMRSEERYELNMAVQOVHANINENANNEEVEWRIEMY 359
QY 361 ISFGIMSLGLSLAVTSPSVSNALNWRSEFSIOSTLGYVALLISTFHVLIYGMKRAFE 420
DB 360 ISFGIMSLGLSLAVTSPSVSNALNWRSEFSIOSTLGYVALLITFHVLIYGMKRAFA 419
QY 421 EBYRFTYPPNFVALVPSIVIL 444
DB 420 EBYRFTYPPNFVALVPSIVIL 443

RESULT 4

Q8CSF0 PRELIMINARY; PRT; 488 AA.
AC Q8CSF0;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE DUDULIN 2 homolog.
GN TSAP6 OR 1010001D0IRIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK078769; BAC37383.1; -.
DR MGD; MGI:1915678; Tsap6.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420 oxidored; 1.
SQ SEQUENCE 488 AA; 54780 MW; FDRC4F42AE503D11 CRC64;

Query Match 56.1%; Score 1318; DB 11; Length 488;
Best Local Similarity 56.2%; Pred. No. 8.4e-98;
Matches 248; Conservative 81; Mismatches 100; Indels 12; Gaps 4;

QY 6 MMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHVIGSRNPKF 65
DB 14 LVDSDSLAE--VP-----KEAPK--VGILSGDFARSLATRLVSGSFVVVGSRNPKR 63
QY 66 ASEFPFHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNMRIN-- 123
DB 64 TAGLFPSLAQVTFQEEAVSSPEVIFVAVFREHYSSLSLADQLAGKILVDVSNPTEKEHL 123
QY 124 QYPSNAEYLAFLPDSLIYKGFNVSAWALQGPKASROVYICSNNIQARQVIELAR 183
DB 124 QHRQNAEYLAFLPACTVVKAFNVISAWALQAGPRDGNROVLICSDQPEAKRTISEMAR 183
QY 184 QLNFIPIDLGLSSAREIENLPLRLFTLWGRPVVAISLATFFFLYSFVRDVIHPYARNQ 243
DB 184 AMGFTPLDMGSLASAREVEAIPRLLPWKVPTLLALGLFVCFYTYNFIKRDVLQPIRKD 243
QY 244 QSDFYKIPIEIVNKTLPVAITLTLVYLAGLLAAAYQLYGTYKRRPPEWLETWLQCRK 303
DB 244 ENKFFYKMPISVNTTTPCVAAYVLLSLVYLPGLVLAALQLRGTGYQRFDPWLDHWLQHRK 303
QY 304 QIGLLSFFFAVHVAYSLCLPMRSEERYELNMAVQOVHANINENANNEEVEWRIEMYISF 363
DB 304 QIGLLSFFFAVHVAYSLCLPLRRSHRYDLVNLAVKQVLANKSRILWEEBWRMEIYLSL 363
QY 364 GIMSGLSLAVTSPSVSNALNWRSEFSIOSTLGYVALLISTFHVLIYGMKRAFE 423
DB 364 GYVALGMLSLAVTSPSIANSINWKEFSFVQSTLGFVALILSTMHTLTYGWTFAFEENH 423

QY 424 YRFYTPNFVALVPSIVIL 444
DB 424 YRFYTPNFVALVPSIVIL 444

RESULT 5

Q99P41 PRELIMINARY; PRT; 488 AA.
AC Q99P41;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Tumor suppressor pHyde.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Dunning; TISSUE=prostatic carcinoma;
RX MEDLINE=20424188; PubMed=10969787;
RA Steiner M.S., Zhang Y., Lu Y.;
RT "Growth inhibition of prostate cancer by an adenovirus expressing a
RT novel tumor suppressor gene, pHyde."
RL Cancer Res. 60:4419-4425(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Dunning; TISSUE=prostatic carcinoma;
RA Lu Y., Rinaldy A.R., Steiner M.S.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Steiner M.S., Rinaldy A.;
RT "A novel prostate cancer associated gene."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF335281; AAK00361.1; -.
DR EMBL; AF238865; AAL78207.1; -.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420 oxidored; 1.
SQ SEQUENCE 488 AA; 54640 MW; 59FF07121919FDCB CRC64;

Query Match 56.0%; Score 1316; DB 11; Length 488;
Best Local Similarity 56.2%; Pred. No. 1.2e-97;
Matches 248; Conservative 81; Mismatches 100; Indels 12; Gaps 4;

QY 6 MMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHVIGSRNPKF 65
DB 14 LVDSDSLAE--VP-----KEAPK--VGILSGDFARSLATRLVSGSFVVVGSRNPKR 63
QY 66 ASEFPFHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNMRIN-- 123
DB 64 TAGLFPSLAQVTFQEEAVSSPEVIFVAVFREHYSSLSLADQLAGKILVDVSNPTEKERL 123
QY 124 QYPSNAEYLAFLPDSLIYKGFNVSAWALQGPKASROVYICSNNIQARQVIELAR 183
DB 124 QHRQNAEYLAFLPACTVVKAFNVISAWALQAGPRDGNROVLICSDQPEAKRTISEMAR 183
QY 184 QLNFIPIDLGLSSAREIENLPLRLFTLWGRPVVAISLATFFFLYSFVRDVIHPYARNQ 243
DB 184 AMGFTPLDMGSLASAREVEAIPRLLPWKVPTLLALGLSTQSYAYNFIKRDVLQPIRKD 243
QY 244 QSDFYKIPIEIVNKTLPVAITLTLVYLAGLLAAAYQLYGTYKRRPPEWLETWLQCRK 303
DB 244 ENKFFYKMPISVNTTTPCVAAYVLLSLVYLPGLVLAALQLRGTGYQRFDPWLDHWLQHRK 303
QY 304 QIGLLSFFFAVHVAYSLCLPMRSEERYELNMAVQOVHANINENANNEEVEWRIEMYISF 363
DB 304 QIGLLSFFFAVHVAYSLCLPLRRSHRYDLVNLAVKQVLANKSRILWEEBWRMEIYLSL 363
QY 364 GIMSGLSLAVTSPSVSNALNWRSEFSIOSTLGYVALLISTFHVLIYGMKRAFE 423
DB 364 GYVALGMLSLAVTSPSIANSINWKEFSFVQSTLGFVALMLSTMHTLTYGWTFAFEENH 423

QY 424 YRFYTPNPNFVLALVLPISIVIL 444
 DB 424 YKFYLPPTFTLTLLPCVITL 444

RESULT 6

O8C159 PRELIMINARY; PRT; 488 AA.
 ID O8C159
 AC O8C159;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Similar to RIKEN CDNA 1010001D01 gene.
 GN TSAP6 OR 1010001D01RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CZECH II;
 RA Strausberg R.;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC037435; AAH37435.1; -.
 DR MGD; MGI:1915678; Tsap6.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR004455; NADPoxred_F420.
 DR Pfam; PF03807; F420 oxidored; 1.
 SQ SEQUENCE 488 AA; 54749 MW; 9A08D99C90CF83F4 CRC64;

Query Match 56.0%; Score 1316; DB 11; Length 488;
 Best Local Similarity 56.2%; Pred. No. 1.2e-97;
 Matches 248; Conservative 81; Mismatches 100; Indels 12; Gaps 4;

QY 6 MMGSPKSLSETCLPENGINGIKDARKVTGVIQSGDFAKSLTIRLCRGYHVIGSRNPKF 65
 DB 14 LVDSDGLAE--VP-----KEAPK--VGILSGDFARSLATRLVSGSGSVVVGSRNPKR 63
 QY 66 ASEFFPHVVDVTHHEDALTNTNIIIFVAIHREHYTSLMDLRHLVKGKILIDVSNMNRIN-- 123
 DB 64 TAGLFFSLAQVTFQEEAVSSPEVIFVAVFREHYSSLCSLADQLAGKILVDVSNPTEKEHL 123
 QY 124 QYPESNAEYLASLFPDSLIVKGFNVVSAMALQLGPKDASROVYICSNNTQARQOVIELAR 183
 DB 124 QHRQSNAEYLASLFPACTVVKAFNVISAMALQAGPRDGNRQVLICSDQPEAKRTISEMAR 183
 QY 184 QINFIPIIDGLSSAREIENLPLRLFTLWKGPVVAISLATFFFLYSFVRDVIHPYARNQ 243
 DB 184 AMGFPLDMGSLASAREVEAIFRLPLPSMKVPTLLALGLFVCFYTYNFIIRDVLPYIRKD 243
 QY 244 QSDFYKIPIEIVNKTLPVIAITLLSLVYLAGLLAAAYQLYYGTKYRRFPPEWLETWLQCRK 303
 DB 244 ENKFKYKPLSVNTTLPVAVYLLSLVYLPGLVLAALQLRRGTYQRFPDWLDHWLQHRK 303
 QY 304 QIGLISFFFAFAMVHVAYSICLPMRSEERYLFNLMAVQOVHANINENSWNEEYWRIMYISF 363
 DB 304 QIGLISFFFAFAMVHVAYSICLPMRSEERYLFNLMAVQOVHANINENSWNEEYWRIMYISL 363
 QY 364 GIMSLGLSLAVTSIPSVSNALNWRKESFIQSTLGYVALLISTFHVLIYGWKRAFEERY 423
 DB 364 GVLALGMLSLAVTSLPSIANSINWKEFSFVQSTLGFVALLISTMTHTLTGWTAFEEENH 423
 QY 424 YRFYTPNPNFVLALVLPISIVIL 444
 DB 424 YKFYLPPTFTLTLLPCVITL 444

RESULT 7
 O80ZF3 PRELIMINARY; PRT; 526 AA.
 ID O80ZF3
 AC O80ZF3;
 DT 01-JUN-2003 (Tremblrel. 24, Created)

DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Tsap6.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM2 ACC331; TISSUE=Bone marrow;
 RX MEDLINE=22506415; PubMed=12606722;
 RA Passer B.J., Nancy-Portebois V., Amzallag N., Prieur S., Cans C.,
 RA Roborel de Climens A., Fiucci G., Boulevard V., Tynnder M., Susini L.,
 RA Morchoisne S.P., Cribble V., Lespagnol A., Dausset J., Oren M.,
 RA Amson R., Telerman A.;
 RT "The p53-inducible TSAP6 gene product regulates apoptosis and the cell
 cycle and interacts with Nix and the Myt1 kinase."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:2284-2289(2003).
 DR EMBL; AY214462; AAC38239.1; -.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR004455; NADPoxred_F420.
 DR Pfam; PF03807; F420 oxidored; 1.
 SQ SEQUENCE 526 AA; 58530 MW; 6306CD717E25200A CRC64;

Query Match 56.0%; Score 1316; DB 11; Length 526;
 Best Local Similarity 56.2%; Pred. No. 1.3e-97;
 Matches 248; Conservative 81; Mismatches 100; Indels 12; Gaps 4;

QY 6 MMGSPKSLSETCLPENGINGIKDARKVTGVIQSGDFAKSLTIRLCRGYHVIGSRNPKF 65
 DB 52 LVDSDGLAE--VP-----KEAPK--VGILSGDFARSLATRLVSGSGSVVVGSRNPKR 101
 QY 66 ASEFFPHVVDVTHHEDALTNTNIIIFVAIHREHYTSLMDLRHLVKGKILIDVSNMNRIN-- 123
 DB 102 TAGLFFSLAQVTFQEEAVSSPEVIFVAVFREHYSSLCSLADQLAGKILVDVSNPTEKEHL 161
 QY 124 QYPESNAEYLASLFPDSLIVKGFNVVSAMALQLGPKDASROVYICSNNTQARQOVIELAR 183
 DB 162 QHRQSNAEYLASLFPACTVVKAFNVISAMALQAGPRDGNRQVLICSDQPEAKRTISEMAR 221
 QY 184 QINFIPIIDGLSSAREIENLPLRLFTLWKGPVVAISLATFFFLYSFVRDVIHPYARNQ 243
 DB 222 AMGFPLDMGSLASAREVEAIFRLPLPSMKVPTLLALGLFVCFYTYNFIIRDVLPYIRKD 281
 QY 244 QSDFYKIPIEIVNKTLPVIAITLLSLVYLAGLLAAAYQLYYGTKYRRFPPEWLETWLQCRK 303
 DB 282 ENKFKYKPLSVNTTLPVAVYLLSLVYLPGLVLAALQLRRGTYQRFPDWLDHWLQHRK 341
 QY 304 QIGLISFFFAFAMVHVAYSICLPMRSEERYLFNLMAVQOVHANINENSWNEEYWRIMYISF 363
 DB 342 QIGLISFFFAFAMVHVAYSICLPMRSEERYLFNLMAVQOVHANINENSWNEEYWRIMYISL 401
 QY 364 GIMSLGLSLAVTSIPSVSNALNWRKESFIQSTLGYVALLISTFHVLIYGWKRAFEERY 423
 DB 402 GVLALGMLSLAVTSLPSIANSINWKEFSFVQSTLGFVALLISTMTHTLTGWTAFEEENH 461
 QY 424 YRFYTPNPNFVLALVLPISIVIL 444
 DB 462 YKFYLPPTFTLTLLPCVITL 482

RESULT 8
 O924Z1 PRELIMINARY; PRT; 514 AA.
 ID O924Z1
 AC O924Z1;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
 DE Dudulin 2.
 GN TSAP6 OR 1010001D01RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Query Match	54.1%	Score 1272;	DB 4;	Length 498;
Best Local Similarity	54.5%;	Pred. No. 4.4e-94;		
Matches 242;	Conservative 82;	Mismatches 108;	Indels 12;	Gaps 4;

QY	3	SISWNGSPKSLSECTCLPNGINGIKDARKVTGVIGSGDPAKSLTIRLIRCGHVIIGSRN	62
Dd	21	SLHLVDSDSLAK--VPD-----BAPK--VGILGSGDFARSLATRLVGSGFKVVGSRN	70
QY	63	PKFASFPFHVVDTTHEDALTKNIIFVAIHREHYTSMDLRHLLVGKILIDYSN--NM	120
Dd	71	PKRTARLFPSAAQVTFQEBAVSSPEYIFVAVFREHYSSLCSLSJDQLAGKILVDVSNPTEQ	130
QY	121	RINQYPESNAEYLASFPPDSLIVGENVVSAMALOLGPXOASROVYICSNNIQAOCVIE	180
Dd	131	EHLQHRESNAEYLAFLPPTCTVWKAFNVISAWTLQAGPRDNKQVPICGDQPEAKRAVSE	190
QY	181	LARQUNFPIDLGSLSSAREIENTPLRFTLMRGFPVVAISAATFFLYSFVRDIHPYA	240
Dd	191	MALANGFMYPDMSGLASAMEVEAMPRLLPAMKVPTLLALGLFVCFYANNFVRDVLOPYV	250
QY	241	RNOQSDEYKPIEBIWNKTLPVIAITLISLVYLAGLLAAAYOLYYGTKYRFPPMLETWLO	300
Dd	251	QESQNKFEKLPEVSVNNTLLPCVAYVLISLVYLLPGVLAALQJLRGTXYQRFPDWLDHWLO	310
QY	301	CRKQJGLSFFFAVHVAYSCLPMRBSERYFLMMAQQOVHANIENTSMNEBEVWRILEMY	360
Dd	311	HRKQJGLSFFCALHALHSFCPLRRRHRYDLVNLAVKQVLANKSHLMVEEBEVWRIMELY	370
QY	361	ISFGIMSLGLSLLAVTSIPSVSNALNMRFSFIQSTLGYVALLISTFHVLIYGWKRAFE	420
Dd	371	LSLGVALLGITUSILA VTSLPSIANSLNMRFSFVQSSLGFWALVLTSLHTLTYGWTRAFE	430
QY	421	EEYRYFYTTPNFVIALVLPISVIL	444
Dd	431	ESRYKFYLPPTFTILLVPCVIL	454

```

RESULT 11
Q9NVB5
ID Q9NVB5 PRELIMINARY; PRT; 488 AA.
AC Q9NVB5;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, last annotation update)
DE Hypothetical protein FLJ10829 (Dudulin 2).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isogai T., Oca T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto D., Wakamatsu A.,
RA Nakamura Y., Nagahari K., Masuko Y., Sasaki N.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Serru V., Manivet P., Lenoir C., Eschwege P., Lambdin D.,
RA Vaubourdoille M., Kellermann O., Loric S.;
RT "Dudulin 2, a new tumor antigen expressed in various human tumors.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK001691; BAA91839.1; -.
DR EMBL; AY029585; AAK50538.1; -.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420_oxidored; I.
DR PROSITE; PS00290; IG_MHC; 1.
KW Hypothetical protein.
SQ SEQUENCE 488 AA; 54616 MW; BC0BCA483335AAD6 CRC64;

```

Query Match	54.0%;	Score 1269;	DB 4;	Length 488;
Best Local Similarity	54.3%;	Pred. No. 7.5e-94;		
Matches 241;	Conservative	83;	Mismatches 108;	Indels 12;
				Gaps 4;

[illegible]

```

RESULT 12
Q8NEW6
ID Q8NEW6 PRELIMINARY; PRT; 487 AA.
AC Q8NEW6;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, last annotation update)
DE PHYde.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lu Y., Beheshti B., Squire J.A., Yang X.J.;
RT "Characterization of a novel apoptosis-inducing gene, hphYde, that
RT inhibits prostate cancer cell growth.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY082673; AAM45136.1; -
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR004455; NADpoxred_F420.
DR Pfam; PF03807; F420_oxidored_I.
DR PROSITE; PS00290; IG_MHC; 1.
SQ SEQUENCE 487 AA; 54501 MW; 80EF95EB55C81EDF CRC64;

```

```

Query Matchn      53.3%; Score 1252.5; DB 4; Length 487;
Best Local Similarity 54.1%; Pred. No. 1.6e-92;
Matches 240; Conservative 83; Mismatches 108; Indels 13; Gaps 5;

QY      3 SISMNGSPKSLSETCLPENGINGIKDKARKVTGVIGSGDFAKSLTIRLIRCGYHVVIGSRN 62
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      11 SLHLVDSDSLAK--VPD-----EAPKVS--ILGSGDFARSLATRLVSGGFKVVGSRN 60

QY      63 PKFASEFFPHVVDVTHHEDALTKNNIIFVAIHREHYTSLMDLRHLLVGGKILIDVSN--NM 120

```



```

Db      61 PKRTARLFPSAAQVTFQEEAVSSPEVIFVAVFREHYSSLCSLSDQLAGKILVDVSNPTEQ 120
QY      121 RINQYPSNAEYLAFLPDSLIVKGFNVSAWALQGPKDASRQYVICSNNIOARQVIE 180
Db      121 EHLQHRRESNAEYLAFLPCTCTVKAFFNVISAWTLQAGPRDGNRQVPICGDQPEAKRAVSE 180
QY      181 LARQNFIPIDLSLSSAREIENLPLRLFTLWGRPVVAISLATFFFLYSFVRDVIHPYA 240
Db      181 MALAMGFMPVDMGSLASAMEVEAMPLRLPAWKVPTLLALGLFVCFYAYNFVRDVLQPYV 240
QY      241 RNQSDPFYKIPIEIVNKTLPVIAITLLSLVYLALAAAYQLYGTYKRRFPPLWETWLO 300
Db      241 QESQNKFFKLPVSVVNTLLPCVAYVLLSLVLPGLAALQLRGTGYQRFPPDMLDHWLO 300
QY      301 CRKQGLLSFFPAMVAVYSLCLPMRSEERYLFNMAYQOVHANIEENSWNEEVRRIEM 360
Db      301 HRKQIGLISFCAALHALYSFCLPLRRARHYDLVNLAVKQVLANKSHLW-VEEVRMEIY 359
QY      361 ISFGIMSLGLSLAVTSIPSVSNALNWRFSFIQSTLGYVALLISTFHVLIYWKRAFE 420
Db      360 LSLGVLAIGTSLSLAVTSIPSIANSLSNWRFSFVQSSLGVALVLSLHTLTYGWTRAFE 419
QY      421 EEEYRFYTPPNEVLAIVLPSIVIL 444
Db      420 ESRYKFYLPPTFTLLVPCVIL 443

```

RESULT 13

```

Q8TF03 PRELIMINARY; PRT; 487 AA.
ID Q8TF03 AC Q8TF03;
RT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Tumor suppressor pHyde.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Steiner M.S., Allay J.A., Wang C.;
RT "A novel prostate-derived tumor suppressor."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF238864; AAL78206.1; -.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420 oxidored; I.
DR PROSITE; PS00290; IG_MHC; 1.
SQ SEQUENCE 487 AA; 54402 MW; AFF16053590E6F68 CRC64;

```

Query Match 53.0%; Score 1245.5; DB 4; Length 487;
 Best Local Similarity 53.8%; Pred. No. 5.8e-92;
 Matches 239; Conservative 83; Mismatches 109; Indels 13; Gaps 5;

```

QY      3 STSMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLIRCGYHVIGSRN 62
Db      11 SLHLVSDSSSLAK--VPD-----EAPKVS--ILGSGDFARSLATRLVSGGFKYVVGSRN 60
QY      63 PKFASEFFPHVVDVTHHEDALTKNIIIVAIHREHYTSLMDLRHLVKGKILIDVSN--NM 120
Db      61 PKRTARLFPSAAQVTFQEEAVSSPEVIFVAVFREHYSSLCSLSDQLAGKILVDVSNPTEQ 120
QY      121 RINQYPSNAEYLAFLPDSLIVKGFNVSAWALQGPKDASRQYVICSNNIOARQVIE 180
Db      121 EHLQHRRESNAEYLAFLPCTCTVKAFFNVISAWTLQAGPRDGNRQVPICGDQPEAKRAVSE 180
QY      181 LARQNFIPIDLSLSSAREIENLPLRLFTLWGRPVVAISLATFFFLYSFVRDVIHPYA 240
Db      181 MALAMGFMPVDMGSLASAMEVEAMPLRLPAWKVPTLLALGLFVCFYAYNFVRDVLQPYV 240

```

```

QY      241 RNQSDPFYKIPIEIVNKTLPVIAITLLSLVYLALAAAYQLYGTYKRRFPPLWETWLO 300
Db      241 QESQNKFFKLPVSVVNTLLPCVAYVLLSLVLPGLAALQLRGTGYQRFPPDMLDHWLO 300
QY      301 CRKQGLLSFFPAMVAVYSLCLPMRSEERYLFNMAYQOVHANIEENSWNEEVRRIEM 360
Db      301 HRKQIGLISFCAALHALYSFCLPLRRARHYDLVNLAVKQVLANKSHLW-VEEVRMEIY 359
QY      361 ISFGIMSLGLSLAVTSIPSVSNALNWRFSFIQSTLGYVALLISTFHVLIYWKRAFE 420
Db      360 LSLGVLAIGTSLSLAVTSIPSIANSLSNWRFSFVQSSLGVALVLSLHTLTYGWTRAFE 419
QY      421 EEEYRFYTPPNEVLAIVLPSIVIL 444
Db      420 ESRYKFYLPPTFTLLVPCVIL 443

```

RESULT 14

```

Q8TDP3 PRELIMINARY; PRT; 456 AA.
ID Q8TDP3 AC Q8TDP3;
RT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE PHYDE II.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang C., Allay J.A., Steiner M.S.;
RT "Second human member of pHyde family, Human pHyde II."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF262322; AAM08128.1; -.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420 oxidored; I.
DR PROSITE; PS00290; IG_MHC; 1.
SQ SEQUENCE 456 AA; 50430 MW; CSF7C7008D55251E CRC64;

```

Query Match 45.7%; Score 1074.5; DB 4; Length 456;
 Best Local Similarity 52.9%; Pred. No. 3.2e-78;
 Matches 209; Conservative 75; Mismatches 98; Indels 13; Gaps 5;

```

QY      3 STSMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLIRCGYHVIGSRN 62
Db      11 SLHLVSDSSSLAK--VPD-----EAPKVS--ILGSGDFARSLATRLVSGGFKYVVGSRN 60
QY      63 PKFASEFFPHVVDVTHHEDALTKNIIIVAIHREHYTSLMDLRHLVKGKILIDVSN--NM 120
Db      61 PKRTARLFPSAAQVTFQEEAVSSPEVIFVAVFREHYSSLCSLSDQLAGKILVDVSNPTEQ 120
QY      121 RINQYPSNAEYLAFLPDSLIVKGFNVSAWALQGPKDASRQYVICSNNIOARQVIE 180
Db      121 EHLQHRRESNAEYLAFLPCTCTVKAFFNVISAWTLQAGPRDGNRQVPICGDQPEAKRAVSE 180
QY      181 LARQNFIPIDLSLSSAREIENLPLRLFTLWGRPVVAISLATFFFLYSFVRDVIHPYA 240
Db      181 MALAMGFMPVDMGSLASAMEVEAMPLRLPAWKVPTLLALGLFVCFYAYNFVRDVLQPYV 240
QY      241 RNQSDPFYKIPIEIVNKTLPVIAITLLSLVYLALAAAYQLYGTYKRRFPPLWETWLO 300
Db      241 QESQNKFFKLPVSVVNTLLPCVAYVLLSLVLPGLAALQLRGTGYQRFPPDMLDHWLO 300
QY      301 CRKQGLLSFFPAMVAVYSLCLPMRSEERYLFNMAYQOVHANIEENSWNEEVRRIEM 360
Db      301 HRKQIGLISFCAALHALYSFCLPLRRARHYDLVNLAVKQVLANKSHLW-VEEVRMEIY 359
QY      361 ISFGIMSLGLSLAVTSIPSVSNALNWRFSFIQ 395
Db      360 LSLGVLAIGTSLSLAVTSIPSIANSLSNWRFSFVQ 394

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM proteoin - protein search, using sw model

Run on: April 16, 2004, 12:33:12 ; Search time 17 Seconds

(without alignments)
1390.579 Million cell updates/sec

Title: US-09-455-486-6

Perfect score: 2351

Sequence: 1 MESISMMSPKSLSETCLPN.....ALVLPSTIVLDLQLCRYPD 454

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	717	30.5	339	1	STEAL_HUMAN
2	133	5.7	223	1	FARE_METJA
3	120.5	5.1	232	1	FARE_METTH
4	111.5	4.7	695	1	FSHR_MACFA
5	109.5	4.7	206	1	Y538_PASMU
6	109	4.6	693	1	FSHR_CHICK
7	109	4.6	712	1	FRE6_YEAST
8	107.5	4.6	629	1	FRE7_YEAST
9	107.5	4.6	695	1	FSHR_HUMAN
10	104.5	4.4	695	1	FSHR_PIG
11	104	4.4	396	1	SOTB_ECOS7
12	104	4.4	396	1	SOTB_ECOS1
13	103	4.4	465	1	NPT1_MOUSE
14	102.5	4.4	391	1	NCCC_NEPOL
15	102	4.3	220	1	Y304_BRUME
16	102	4.3	220	1	Y301_BRUST
17	100	4.3	476	1	RF3_SACBA
18	99.5	4.2	452	1	NU4M_BRALA
19	99	4.2	365	1	SPR1_HUMAN
20	98	4.2	3411	1	POLG_YEFV1
21	98	4.2	3411	1	POLG_YEFV2
22	97.5	4.1	268	1	YD49_AQUAE
23	97.5	4.1	695	1	FSHR_BOVIN
24	97	4.1	218	1	YG46_XANAC
25	97	4.1	692	1	FSHR_RAT
26	96.5	4.1	686	1	FRE1_YEAST
27	96	4.1	261	1	PHSC_ECOLI
28	95	4.0	202	1	YAT1_PSEAE
29	95	4.0	393	1	NUCC_ARATH
30	95	4.0	393	1	NUCC_SPIOL
31	95	4.0	692	1	FSHR_MOUSE
32	94.5	4.0	492	1	SECY_CYAPA
33	94.5	4.0	695	1	FSHR_SHEEP

34	93.5	4.0	365	1	SPR1_MOUSE	Q8bfq3 mus musculus
35	93.5	4.0	452	1	NU4M_BRALA	Q47423 branchiosto
36	93.5	4.0	499	1	GSNR_PLAF7	Q15770 plasmodium
37	93.5	4.0	530	1	AAAI_MOUSE	Q9jmh8 mus musculus
38	93.5	4.0	1233	1	VLI_FEOVD	P15024 reovirus (t
39	93	4.0	388	1	HMC3_DESVH	P33390 desulfovibr
40	93	4.0	459	1	NU4M_BOVIN	P03910 bos taurus
41	93	4.0	521	1	YT25_CAEBL	Q10934 caenorhabdi
42	92.5	3.9	490	1	GTR1_CHICK	P46896 gallus gall
43	92.5	3.9	500	1	NU4C_OENHO	P58419 oenothera h
44	92.5	3.9	883	1	YHL6_YEAST	P38781 saccharomyc
45	92.5	3.9	1769	1	YUK9_YEAST	P42945 saccharomyc

ALIGNMENTS

RESULT 1

STEAL_HUMAN
ID STEAL_HUMAN STANDARD: PRT; 339 AA.

AC Q9UHE8; Q95034; 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Six transmembrane epithelial antigen of prostate.
GN STEAP OR STEAP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20056277; PubMed=10588738;
RA Hubert R.S., Vivanco I., Chen E., Rastegar S., Jeong K.,
RA Mitchell S.C., Madraswala R., Zhou Y., Kuo J., Raitano A.B.,
RA Jakobovits A., Saffran D.C., Afar D.E.H.;
RT "STEAP: a prostate-specific cell-surface antigen highly expressed in
RT human prostate tumors.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:14523-14528(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Abu-Threideh J., Stoneking T., Langston Y., Maupin R.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skalek U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- TISSUE SPECIFICITY: Highly expressed in prostate tumors.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF186249; AAF17479.1; -.
DR EMBL; AC005053; AAC79150.1; ALT_INIT.
DR EMBL; AC004969; AAD15620.2; -.
DR EMBL; BC011802; AAH1802.1; -.
DR Genew; HGNC:11378; STEAP.
DR MIM; 604415; -.
DR GO; GO:0005887; C:integral to plasma membrane; TAS.
DR GO; GO:0005911; C:intercellular junction; TAS.
DR GO; GO:0015267; F:channel/pore class transporter activity; TAS.
KW Transmembrane; Antigen.
FT TRANSMEM 71 91 POTENTIAL.
FT TRANSMEM 119 139 POTENTIAL.
FT TRANSMEM 164 184 POTENTIAL.
FT TRANSMEM 218 238 POTENTIAL.
FT TRANSMEM 258 278 POTENTIAL.
FT TRANSMEM 291 311 POTENTIAL.
SQ SEQUENCE 339 AA; 39851 MW; 55443A170C870387 CRC64;

Query Match 30.5%; Score 717; DB 1; Length 339;
Best Local Similarity 54.9%; Pred. No. 6e-48;
Matches 130; Conservative 48; Mismatches 59; Indels 0; Gaps 0;

QY 208 LFTLMRGPVVAISLATPFFPLYSFVRDVIHPYARNOQSDFYKIPBIVKTLPIVAITLL 267
DB 67 LFPQWHLPIKIAATIASLTLTYLLREVHPLATSHQOYFYKIPILVINKVLPMSITLL 126
QY 268 SLVYLLAGLLAAAYQLYGTYKRRPFWLETLQCKRQLGLSFPFAMVHYAVSLCLPMRR 327
DB 127 ALVYLPGVTAIVOLHNGTKYKPPHMLDKMMLTRKQGLLSFPFAVLHAYISLSYPMRR 186
QY 328 SERYFLNMAYYQOVHANINSEWNEEVRIMYISFGIMSLGLSLAVTSIPSVSNALN 387
DB 187 SYRYKLNWAYYYQVQNKEDAWIEHDWRMEIYVSLGIVGLAIALAVTSIPSVSDSLT 246
QY 388 WRFFSFIQSTLGVALLISTFHVLIYGWKRAFEERYRPTPNFVALVPSIVIL 444
DB 247 WRFFHYIQSKLGIIVSLTLGTHALIFAMNKWIDIKQFVWYTPPTFMAVFLPIVILI 303

RESULT 2
F4RE_METJA STANDARD; PRT; 223 AA.
ID F4RE_METJA
AC Q58896;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative F420-dependent NADP reductase (EC 1.-.-.-).
GN MJ1501.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
CX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrman J.L., Nguyen D.,
RA Uutterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotten M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073(1996).
CC -!- FUNCTION: Catalyzes the reduction of F420 with NADP(+) and the
```

```
CC reduction of NADP(+) with F420H(2). Probably couples the NADP-
CC dependent oxidation of the alcohol to the aldehyde with the F420-
CC dependent reduction of Co(2) to methane (anabolic function) (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: NADPH + coenzyme F420 = NADP(+) + reduced
CC coenzyme F420.
CC -!- SIMILARITY: TO M.ORGANOPHILUM F420-DEPENDENT NADP REDUCTASE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67591; AAB99514.1; -.
DR PIR; D64487; D64487.
DR TIGR; MJ1501; -.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420_oxidored; 1.
DR TIGRFAMs; TIGR00301; TIGR00301; 1.
DR Hypothetical protein; Oxidoreductase; NADP; Complete proteome.
KW SEQUENCE 223 AA; 24068 MW; 2370BBD6F5BD3D9 CRC64;

Query Match 5.7%; Score 133; DB 1; Length 223;
Best Local Similarity 21.9%; Pred. No. 0.003;
Matches 46; Conservative 52; Mismatches 78; Indels 34; Gaps 7;

QY 37 GSGDPAKSLTIRLRGCVHVIIGSRNPKFASFPFHVVDV-----THEDALT 84
DB 8 GTBDQGFGLRLRAK-NKKIILIGSRKKEKAEAKAKEILKRGIEADIIGLENKDAK 66
QY 85 KTNIIFVAIHREH-YTSLMDLRHLVGLIDV-----SNNMRINQPE-SNAEYL 133
DB 67 EGDVYILSLPYEYTLSTIKQLKELKGIIVSGVPLATAIGDKPTRLFPDGSVAEMV 126
QY 134 ASLFPDSLIVKGFNVSAWALQLGPKDASROYICSNNIQARQVIELARQINFI-PIDL 192
DB 127 QNVLEKSKVSAFQNVCHAVLELDNFPVDCDILVCGNDEAKKVVIDLANQIDGVRAIDC 186
QY 193 GSLSSAREIENLPLRLFTLMRGFVVAISL 222
DB 187 GNLEKSRRIIEAIT-----PLIIGLNI 207

RESULT 3
F4RE_METTH STANDARD; PRT; 232 AA.
ID F4RE_METTH
AC O26350;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative F420-dependent NADP reductase (EC 1.-.-.-).
GN MTH248.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
CX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakey D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Utwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RA "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
```



```
CC -I- FUNCTION: Catalyzes the reduction of F420 with NADP(+) and the  
CC reduction of NADP(+) with F420H(2). Probably couples the NADP-  
CC dependent oxidation of the alcohol to the aldehyde with the F420-  
CC dependent reduction of CO(2) to methane (anabolic function) (By  
CC similarity).
```

```
CC CC CATALYTIC ACTIVITY: NADPH + coenzyme F420 = NADP(+) + reduced  
CC coenzyme F420.  
CC CC SIMILARITY: TO M.ORGANOPHYLUM F420-DEPENDENT NADP REDUCTASE.
```

```
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).
```

```
DR DR EMBL; AE00811; AAB84754.1; -.  
DR PIR; A69131; A69131.  
DR InterPro; IPR004455; NADPOxred_F420.  
DR Pfam; PF03807; F420_Oxidored; 1.  
DR TIGRFAMs; TIGR00301; TIGR00301; 1.  
KM Hypothetical protein; Oxidoreductase; NADP; Complete proteome.  
SQ SEQUENCE 232 AA; 24539 MW; AICE6OABC8474296 CRC64;
```

```
Query Match          5.1%; Score 120.5; DB 1; Length 232;  
Best Local Similarity 23.8%; Pred. No. 0.029;  
Matches 49; Conservative 42; Mismatches 90; Indels 25; Gaps 6;
```

```
QY      30 KVTGCVTGSGDFAKSLITRLRCGYHVVIGSRNPKFASFPFHVDTHNE----- 80  
        ||::|||::|||::|||::|||::|||::|||::|||:  
Db       8 KIAV-IGGTGDQGGLALRFVAAGEEVITGSRDAAEKASKASKYLEIAGRDDISVEGATN 66  
  
QY      81 -DALTKNNIIFFAIH-REHYTSLEMDRLHLVGKILIDVS-----NMNRINQPESN 129  
        ||::|||::|||::|||::|||::|||::|||:  
Db     67 PDAAASADVVLTVPLQAQMVTLASIRDQRDKVLIDATVPIDSCIGSAVRYYDLWEGS 126  
  
QY      130 AEYLASF--PDSLIVKGFNVSAMWLQLGPKDASRVVTCGNNIQARQQVIELARQLN 186  
        ||::|||::|||::|||::|||::|||::|||:  
Db     127 AAEARAFLREQGRVAAAFFNNIGASALLEVSEPVDCCLVASDHRALEVAAEIAEKID 186  
  
QY      187 FI-PIDLGLSSAREIENTLPRLFTL 211  
        :||::|||::|||::|||::|||:  
Db     187 GVRAIECGGLENNARIITEKITPLLNL 212
```

```
Dd
```

```
RESULT 4  
FSHR_MACFA STANDARD; PRT; 695 AA.  
ID   FSHR_MACFA  
AC   P32212;  
DT   01-OCT-1993 (Rel. 27, Created)  
DI   01-OCT-1993 (Rel. 27, Last sequence update)  
DE   28-FEB-2003 (Rel. 41, Last annotation update)  
DE   Follicle stimulating hormone receptor precursor (FSH-R) (Follitropin  
    receptor).  
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC   Cercopithecinae; Macaca.  
OX   NCBI_TaxID=9541;  
RN   [1]  
RP   SEQUENCE FROM N.A.  
RC   TISSUE=Testis;  
RX   MEDLINE=94071854; PubMed=7504463;  
RA   Gromoli U., Dankbar B., Sharma R.S., Nieschlag E.;  
RT   "Molecular cloning of the testicular follicle stimulating hormone  
    receptor of the non human primate Macaca fascicularis and  
    identification of multiple transcripts in the testis."  
RT   Biochem. Biophys. Res. Commun. 196;1066-1072(1993).  
CC   -I- FUNCTION: Receptor for follicle stimulating hormone. The activity  
    of this receptor is mediated by G proteins which activate  
    adenylyate cyclase.
```

```
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC FSH/LSH/TSH subfamily.
CC -1- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X74454; CAAS2463.1; -.
DR PIR; JN0898; JN0898.
DR HSSP; P23945; 1XUN.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR001611; LRR_Nterm.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR_3.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00237; GPCRRHODPSN.
DR SMART; SM00013; LRRT; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KM G-protein coupled receptor; Transmembrane; glycoprotein; signal;
FM phosphorylation; Repeat; leucine-rich repeat.
KT SIGNAL.
FT CHAIN 1 17 POTENTIAL.
FT DOMAIN 18 695 POLICICLE STIMULATING HORMONE RECEPTOR.
FT DOMAIN 18 366 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 367 387 1 (POTENTIAL).
FT DOMAIN 388 398 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 399 421 2 (POTENTIAL).
FT DOMAIN 422 443 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 444 465 3 (POTENTIAL).
FT DOMAIN 466 485 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 486 508 4 (POTENTIAL).
FT DOMAIN 509 528 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 529 550 5 (POTENTIAL).
FT DOMAIN 551 573 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 574 597 6 (POTENTIAL).
FT DOMAIN 598 608 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 609 630 7 (POTENTIAL).
FT DOMAIN 631 695 CYTOPLASMIC (POTENTIAL).
FT REPEAT 44 68 LRR 1.
FT REPEAT 69 93 LRR 2.
FT REPEAT 119 143 LRR 3.
FT REPEAT 170 192 LRR 4.
FT REPEAT 193 216 LRR 5.
FT REPEAT 218 240 LRR 6.
FT DISULFID 442 517 BY SIMILARITY.
FT CARBOHYD 191 191 N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT CARBOHYD 199 199 N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT CARBOHYD 318 318 N-LINKED (GLCNAC. . . ) (POTENTIAL).
SQ SEQUENCE 695 AA; 78343 MW; 0D60A233729B5250 CRC64;

Query Match 4.7%; Score 111.5; DB 1; Length 695;
Best Local Similarity 17.0%; Pred. No. 0.56;
Matches 84; Conservative 75; Mismatches 151; Indels 183; Gaps 16;
```

```

QY 242 NQOSDFYKIPFIVNKTLPVIAITLLSLVY----- 271
DB 247 RSTYNLKKLP-----SLEKLVALMEASLTYPESHCCAFANMRROISELHPICNKSILRQEV 301
QY 272 -----LAGLLAAAYQLYYGTXYRRF-----PPWLETWLQCRKQLG 306
DB 302 DYMOTRGQRSSLAEDNESSYSRGFDMTYAEFDYDLQNEVDVTCSEKPDAPNCPCEDILG 361
QY 307 -----LISFFPAMVYA-----YSLCLPMRRSERYFLNMAVQOVHANIENS 348
DB 362 YNLRVLWIFISILATGNIIIVLTLTTSQYKLYP-----RFLMCNLAAPAD----- 408
QY 349 WNEEEVWRIEMYISFGINSGLSLAVTSIPSVSNALNRE-----FSFIQSTLG 399
DB 409 -----LCIGTYLLLIASVDIHTKSQYHNYAIDWQTGAGCDAGFFTFASELS 456
QY 400 ---YVALLISTFH-----VLIYGMKRAFEERYRYTPPNFV-LAL 436
DB 457 VYTLTAITLERWHTITHANQLDCKVHRHAASVVMGWIFAFAAALFPIFGISSYMKVSI 516
QY 437 VLPSIVLDLQL 449
DB 517 CLPMDIDSPLSQL 529

RESULT 5
Y538_PASMTU STANDARD; PRT; 206 AA.
ID Y538_PASMTU STANDARD; PRT; 206 AA.
AC O9CN97;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein PM0538.
GN PM0538.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.U., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: Belongs to the UPF0191 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; AE006089; AK02622.1; -
CC DR HAMAP; MF_01207; -; 1.
CC DR InterPro; IPR007916; UPF0191.
CC DR Pfam; PF05252; UPF0191; 1.
CC KW Hypothetical protein; Transmembrane; Complete proteome.
CC FT TRANSMEM 7 POTENTIAL.
CC FT TRANSMEM 41 63 POTENTIAL.
CC FT TRANSMEM 76 98 POTENTIAL.
CC FT TRANSMEM 113 135 POTENTIAL.
CC FT TRANSMEM 142 164 POTENTIAL.
CC FT TRANSMEM 169 191 POTENTIAL.
CC SQ SEQUENCE 206 AA; 23653 MW; 2F277729D7F11194 CRC64;

Query Match 4.7%; Score 109.5; DB 1; Length 206;
Best local Similarity 23.8%; Pred. No. 0.18;
Matches 56; Conservative 44; Mismatches 72; Indels 63; Gaps 13;

```

```

QY 228 LYSFVRDVH-----PYA-----RNQOSDFYKIPFIVNKTLPVIAITLLSLVYLAGL 275
DB 1 MLSLFRILIHVCCUGFVAMLAWLLSGDESQLGADPIKEIQHFLGSAITILLIMFILG- 59
QY 276 LAAAYQLYYGTXYRRFPWLETWLQCRKQLGLSFFPAMVHVAAYSLCLPMRRSERYFLN 335
DB 60 -----KVFFYLK-----QPOLQV---LRRALGIWAMFYVLHV-YA-----YLALE 96
QY 336 MAYQOVHANIENSWNEEEVWRIEMYISFGINSGLSLAVTSIPSVSNALNREFSFIQ 395
DB 97 LGYD-----FSLFVQELVNR--GYLLIGAIAFLITLMALSSWSYLR--LKMGMWFYL 146
QY 396 STLGYVALLISTFHYLIYGMKRAFEERYRYTPPNFVALVLPISIVLDLQLC 450
DB 147 HOLGYVALLLGAIR---YVW-----SVKNVTFSSML--YLLISIMILC 184

RESULT 6
FSHR_CHICK STANDARD; PRT; 693 AA.
ID FSHR_CHICK STANDARD; PRT; 693 AA.
AC P79763; Q90719;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Follicle stimulating hormone receptor precursor (FSH-R) (Folliotropin
DE receptor).
GN FSHR.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=97473503; PubMed=9332357;
RA Wakabayashi N., Suzuki A., Hoshino H., Nishimori K., Mizuno S.;
RT "The cDNA cloning and transient expression of a chicken gene encoding
RT a follicle-stimulating hormone receptor.";
RL Gene 197:121-127(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=White leghorn;
RX MEDLINE=97057887; PubMed=8902217;
RA You S., Bridgham J.T., Foster D.N., Johnson A.L.;
RT "Characterization of the chicken follicle-stimulating hormone
RT receptor (cFSH-R) complementary deoxyribonucleic acid, and expression
RT of cFSH-R messenger ribonucleic acid in the ovary.";
RL Biol. Reprod. 55:1055-1062(1996).
CC -1- FUNCTION: Receptor for follicle stimulating hormone. The activity
CC of this receptor is mediated by G proteins which activate
CC adenylyl cyclase.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC FSH/LSH/TSH subfamily.
CC -1- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; D87871; BA013487.1; -
CC DR EMBL; U51097; AAC60030.1; -
CC DR HSSP; P23945; 1XUN.
CC DR InterPro; IPR00276; GPCR_Rhodpsn.
CC DR InterPro; IPR001611; LRR.
CC DR InterPro; IPR000372; LRR_Nterm.
CC Pfam; PF00001; 7tm_1; 1.

```



```

DR Pfam; P000560; LRR; 2.
DR PRINTS; PR00237; GPCRHRDOPSN.
DR SMART; SMO0013; LRRT; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein; signal;
KW phosphorylation; Repeat; leucine-rich repeat.
FT SIGNAL 1 17
FT CHAIN 18 693
FT DOMAIN 18 366
FT TRANSMEM 367 387
FT DOMAIN 388 398
FT TRANSMEM 399 421
FT DOMAIN 422 443
FT TRANSMEM 444 465
FT DOMAIN 466 485
FT TRANSMEM 486 508
FT DOMAIN 509 528
FT TRANSMEM 529 550
FT DOMAIN 551 573
FT TRANSMEM 574 597
FT DOMAIN 598 608
FT TRANSMEM 609 630
FT DOMAIN 631 693
FT REPEAT 45 68
FT REPEAT 69 93
FT REPEAT 95 118
FT REPEAT 119 143
FT REPEAT 168 192
FT REPEAT 193 216
FT REPEAT 218 240
FT DISULFID 442 517
FT CARBOHYD 47 47
FT CARBOHYD 191 191
FT CARBOHYD 199 199
FT CARBOHYD 268 268
FT CONFLICT 4 4
FT CONFLICT 88 88
FT CONFLICT 140 140
FT CONFLICT 174 174
FT CONFLICT 191 191
FT CONFLICT 329 329
SQ SEQUENCE 693 AA; 78697 MW; 46F98699635A1BEC CRC64;

Query Match 4.6%; Score 109; DB 1; Length 693;
Best Local Similarity 17.2%; Pred. No. 0.87;
Matches 100; Conservative 72; Mismatches 173; Indels 238; Gaps 21;

QY 18 LKNG-INGIKDARKVTG-----VIGSGDFAKSLTIRLCGYHVVIGSRNPKFASEFFP 71
DB 61 IPKGAFTGLHDLKEIKESQNDALIEIGNVFSSL-----PKL----- 97
QY 72 HVDVTHHEDALTKNTIFVAIHREHYTSLMDLRHLV-----KILI 114
DB 98 -----HEIRIEKANKL-MKIDQDAFQHLPSLRYLISNTGLSFLPVVHKVHSFQKVL 149
QY 115 DVSNNNRINQYPESNAEYLAFLPDSLIVKGFNVSAWALQLGPKDASROVYICSNNTIOA 174
DB 150 DVQDNIHIRTIERNTFMGLSS-----ESVILR-----LNKNGIOE 184
QY 175 RQQVIELARQINFIPIDLGSLSSAREIENLPRLFTLMRGVVAISLATFFFLYSFVRD 234
DB 185 IKD-----HAFNGTCLDELNLSNLYLEKPEKVFQGAIGVVLDISRTRISFLSHGLE 239
QY 235 VHPYARNQSDFYKIPIEIVNKLPIVAITLLSLVYLAGL----- 275
DB 240 FIKKLRASTYKLLKLP--DVNKRSLIEANFTYPSHCAPTNRKTONTFYPISMSPA 297
QY 276 -----LAAAYQLYGYGTY-----RRF-----PWLFTWLQCR 302
DB 298 KQDLGQGTGRKGRRSAAEDYISHYGTGFGPVENEFYGLCNEVVDFVCSFKPDAPNCE 357
QY 303 KQLG-----LSFFFAVHVA-----YSLCLPMRBSERYLFLNMAYQOVHAN 344

```

```

DB 358 DIMGNVLRVLIMFINILAITGNTVLLIILISSQKLTVP-----RFLMCNLAFAAD----- 408
QY 345 IENSMNEEBWRIEWYISFGIMSLGLSLAVTSIPSVSNALNWR-----FSFIQ 395
DB 409 -----LCIGIYLLFIASVDIQTKSRYYNVAIDWQGAGCNAGFTVFA 452
QY 396 STLGVALLISTFH-----VLIYGWKRAF----- 419
DB 453 SELSVYTLVTILERHITTYAMQNRKYLRAVILIMFGMMFAFTVALLPIFGISSYM 512
QY 420 -----EEBYRFPYTPNFVALVLPSTVILDLQLCRY 452
DB 513 KVICLPMHIETPFQAYV--IFLVLVNLAFAVILICICICICY 553

RESULT 7
FREQ_YEAST STANDARD; PRT; 712 AA.
ID FREQ_YEAST
AC Q12473;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ferric reductase transmembrane component 6 precursor (EC 1.16.1.7)
DE (Ferric-chelate reductase 6).
GN FREQ OR YL051C OR L0593.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_Taxid=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=97313267; PubMed=9169871;
RA Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansorge W.,
RA Benes V., Brueckner M., Delius H., Dubois E., Duesterhoeft A.,
RA Eutian K.-D., Floeth H., Hilbert H., Hebling U., Heumann K.,
RA Heuss-Neitzel D., Hilbert H., Hilger F., Kietz K., Koetter P.,
RA Louis E.J., Messenguy F., Mewes H.-W., Miesga T., Moestl D.,
RA Mueller-Auer S., Neutwich U., Obermaier B., Piravandi E., Pohl T.M.,
RA Portetelle D., Purnelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
RA Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz S.,
RA Underwood A.P., Urrestarazu L.A., Vandenbol M., Verhasselt P.,
RA Viereendeels F., Voet M., Volckaert G., Voss H., Wambut R., Wedler E.,
RA Medler H., Zimmermann F.K., Zollner A., Hand J., Hohnsels J.D.;
RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RL Nature 387:87-90(1997).
CC -1- CATALYTIC ACTIVITY: NADH + 2 Fe(3+) = NAD(+) + 2 Fe(2+).
CC -1- COFACTOR: FAD (Probable).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: Belongs to the FRE / CYB8 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; Z47973; CAA88006.1; -
DR EMBL; Z73156; CAA97503.1; -
DR PIR; S50969; S50969.
DR Germonline; 142046; -.
DR SGD; S0003974; FREQ.
DR InterPro; IPR002916; Ferric reduct.
DR Pfam; PF01794; Ferric_reduct; 1.
KW Oxidoreductase; Electron transport; Transmembrane; Iron transport;
KW FAD; NAD; Glycoprotein; signal; Multigene family.
FT SIGNAL 1 17
FT CHAIN 18 712
FT NP_BIND 493 499 FAD (Potential).

```

```

FT TRANSMEM 168 188 POTENTIAL.
FT TRANSMEM 245 265 POTENTIAL.
FT TRANSMEM 288 308 POTENTIAL.
FT TRANSMEM 329 349 POTENTIAL.
FT TRANSMEM 361 381 POTENTIAL.
FT TRANSMEM 388 408 POTENTIAL.
FT TRANSMEM 417 437 POTENTIAL.
FT TRANSMEM 478 498 POTENTIAL.
FT TRANSMEM 550 570 POTENTIAL.
FT DOMAIN 553 556 POLY-LEU.
FT CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 450 450 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 655 655 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 712 AA; 81989 MW; 5224F12B51544BAA CRC64;

```

Query Match 4.6%; Score 109; DB 1; Length 712;
 Best Local Similarity 18.8%; Pred. No. 0.9;
 Matches 72; Conservative 51; Mismatches 119; Indels 142; Gaps 17;

```

QY 88 IIFVAIHREHYTSL-----WDLRLVLGKILIDVSNMNRINQYPSNAEYLAST 136
DB 183 IIAVFHMSHYNGLRALFASRFVNYIRGHFVLPFLVD---KHANHKFLNVEVFTGL 238
QY 137 FPDSLIVGFNVSAWALQLPKASRQVYICSNNIQARQOVIELARQINFIPIDGLSL 196
DB 239 MPNSL-----EAW----- 246
QY 197 SAREIENLPLRLFTLMRGFVVAISLATFFFL-YSFVRDVIHPYARNOQSDPYKIPFIV 255
DB 247 -----IIFGYTLANIIFLSISYIIDPYNLIENSHLSQFTRL--LA 284
QY 256 NKTLPVIAITLTLVLYLAGLLAAAYQLYGTQYRRFPFMLETLQCRKQGLLSFFPAMV 315
DB 285 DRS-GILAFITQPLIIFTARNSFLEFLTGKFNSE-----ISFHKWIGRIMVLNATI 336
QY 316 H-VAVSLCLPMRSEKRYFLNMAVYQOVHANIEENSWNEEVRIEYISFGIMSLGLSL 374
DB 337 HSLSYSL-----FAIINHAFK-----ISNK-----QLYWKFGIASITVLCVL 373
QY 375 AVTSIPSVSNALNWRSEFISQTLGYVALLI--STFHVLIY-GWKRAFE-----EE 422
DB 374 LVLSLGIVRK---RHYEFPLVTHIILALLFFYCCWQHVKIFNGWKEMIVVSLIWGLEK 429
QY 423 YRFPYTPPNFVALVLPISIVLDL 446
DB 430 LFRIMW---ILQFRFPKATLINTL 449

```

RESULT 8
 FRET YEAST STANDARD; PRT; 629 AA.
 AC Q12333;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ferric reductase transmembrane component 7 (EC 1.16.1.7) (Ferric-chelate reductase 7).
 GN FRET OR YOL152W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / FY1679;
 RX MEDLINE=96132030; PubMed=8553699;
 RA Casamayor A., Aldea M., Casas C., Herrero E., Gamo F.J.,
 RA Lafuente M.J., Gancedo C., Arino J.;
 RT "DNA sequence analysis of a 13 kbp fragment of the left arm of yeast
 RT chromosome XV containing seven new open reading frames.";
 RL Yeast 11:1281-1288(1995).

```

CC -1- CATALYTIC ACTIVITY: NADH + 2 Fe(3+) = NAD(+) + 2 Fe(2+).
CC -1- COFACTOR: FAD (Probable).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: Belongs to the FRET / CYB family.

```

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC EMBL; Z48239; CAA88276.1; -
CC EMBL; Z74894; CAA99174.1; -
CC PIR; S60385; S60385.
CC Germonline; 143574; -.
DR SGD; S0005512; FRET.
DR InterPro; IPR002916; Ferric_reduct.
DR Pfam; PF01794; Ferric_reduct; 1.
KW Oxidoreductase; Electron transport; Transmembrane; Iron transport;
KW FAD; NAD; Glycoprotein; Multigene family.
FT NP BIND 369 375 FAD (POTENTIAL).
FT TRANSMEM 46 66 POTENTIAL.
FT TRANSMEM 108 128 POTENTIAL.
FT TRANSMEM 168 188 POTENTIAL.
FT TRANSMEM 195 215 POTENTIAL.
FT TRANSMEM 238 258 POTENTIAL.
FT TRANSMEM 266 286 POTENTIAL.
FT TRANSMEM 293 313 POTENTIAL.
FT TRANSMEM 422 442 POTENTIAL.
FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 541 541 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 629 AA; 71996 MW; 2384480B9289C16F CRC64;

```

Query Match 4.6%; Score 107.5; DB 1; Length 629;
 Best Local Similarity 21.7%; Pred. No. 1;
 Matches 57; Conservative 46; Mismatches 77; Indels 83; Gaps 17;

```

QY 214 GPVVVAISLATFFFLYSFVRDVIHPYARNOQSDPYKIPFIVNKTLPVIAITLTLVY-L 272
DB 123 GTFVLVMAATLTLYLCVFP--HPFYR-PCAGFGSPPLSV--RAGIMASLVPFVSL 175
QY 273 AG-----LAAAYQLYGTQYRRFPFMLETLQCRKQGLLSFFPAMVAVXSLCLPM 325
DB 176 SGKINIVGLVGLSYE-----KINITHQW-----ASILCLFFSWVHV-----IPF 215
QY 326 RRSERYFLNMAVYQOVHANIEENSWNEEVRIEYISFGIMSLGLSLAVTSIP----- 380
DB 216 LRQARH--EGYERKH-----QRWASDMMR-----SGVPPILFLNLMLSSLPARRH 262
QY 381 --SVSNALNWRSEFISQTLGYVALLISTFHV-----LTYGKRAFESEYYR- 425
DB 263 FYRIFLQLHW---ILAVGFYISLF--YHYVPELNSHMYLVATIVW---FAQLFYRL 311
QY 426 ---FYTPPNFVALVLPISIVIL 444
DB 312 AVKGYLRPGRSFMASTIANVISIV 334

```

RESULT 9
 FSHR HUMAN STANDARD; PRT; 695 AA.
 AC P23945; Q16225;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Follicle stimulating hormone receptor precursor (FSH-R) (Follltropin receptor).
 GN FSHR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;
 RN [1]
 RA SEQUENCE FROM N.A., AND VARIANTS ALA-307 AND SER-680.
 RC TISSUE=Ovary;
 RX MEDLINE=91222171; PubMed=1709010;
 RA Minegishi T., Nakamura K., Takakura Y., Ibuki Y., Igarashi M.;
 RT "Cloning and sequencing of human FSH receptor cDNA.";
 RL Biochem. Biophys. Res. Commun. 175:1125-1130(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=93246012; PubMed=1301382;
 RA Kelton C.A., Cheng S.V., Nugent N.P., Schweickhardt R.L.,
 RA Rosenthal J.L., Overton S.A., Wands G.D., Kuzeja J.B., Luchette C.A.,
 RA Chappel S.C.;
 RT "The cloning of the human follicle stimulating hormone receptor and
 RT its expression in COS-7, CHO, and Y-1 cells.";
 RL Mol. Cell. Endocrinol. 89:141-151(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Tilly L.T., Aihara T., Nishimori K., Jai X.-C., Billig H.,
 RA Kowalski K.I., Perlas E.A., Heueh A.J.;
 RL Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 1-342 FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=93075197; PubMed=1359889;
 RA Gromoll J., Gudermann T., Nieschlag E.;
 RT "Molecular cloning of a truncated isoform of the human follicle
 RT stimulating hormone receptor.";
 RL Biochem. Biophys. Res. Commun. 188:1077-1083(1992).
 RN [5]
 RP SEQUENCE OF 1-51 FROM N.A.
 RX MEDLINE=95011044; PubMed=7926278;
 RA Gromoll J., Danbar B., Gudermann T.;
 RT "Characterization of the 5' flanking region of the human follicle-
 RT stimulating hormone receptor gene.";
 RL Mol. Cell. Endocrinol. 102:93-102(1994).
 RN [6]
 RP SEQUENCE OF 286-695 FROM N.A., AND VARIANT SER-680.
 RX MEDLINE=9500244; PubMed=7916967;
 RA Gromoll J., Ried T., Holtgreve-Grez H., Nieschlag E., Gudermann T.;
 RT "Localization of the human FSH receptor to chromosome 2p21 using a
 RT genomic probe comprising exon 10.";
 RL J. Mol. Endocrinol. 12:265-271(1994).
 RN [7]
 RP 3D-STRUCTURE MODELING OF 49-228.
 RX MEDLINE=96363672; PubMed=8747461;
 RA Jiang X., Dreano M., Buckler D.R., Cheng S., Ythier A., Wu H.,
 RA Hendrickson W.A., el Tayar N.;
 RT "Structural predictions for the ligand-binding region of glycoprotein
 RT hormone receptors and the nature of hormone-receptor interactions.";
 RL Structure 3:1341-1353(1995).
 RN [8]
 RP VARIANTS ALA-307; ARG-524 AND SER-680.
 RX MEDLINE=99318093; PubMed=10391209;
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipschutz R., Daley G.Q.,
 RA Lander E.S.;
 RT "Characterization of single-nucleotide polymorphisms in coding regions
 RT of human genes.";
 RL Nat. Genet. 22:231-238(1999).
 RN [9]
 RP ERRATUM.
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipschutz R., Daley G.Q.,
 RA Lander E.S.;
 RL Nat. Genet. 23:373-373(1999).
 RN [10]
 RP VARIANTS ALA-307 AND SER-680.
 RX MEDLINE=22054685; PubMed=12059813;

RA Asatiani K., Gromoll J., Eckardstein S.V., Zitzmann M., Nieschlag E.,
 RA Simoni M.;
 RT "Distribution and function of FSH receptor genetic variants in normal
 RT men.";
 RL Andrologia 34:172-176(2002).
 RN [11]
 RP VARIANT OHSS ILE-449.
 RX MEDLINE=22812037; PubMed=12930927;
 RA Vasseur C., Rodien P., Beau I., Desroches A., Gerard C.,
 RA de Poncheville L., Chaplot S., Savagner F., Crove A., Mathieu E.,
 RA Lahlou N., Descamps P., Mitrani M.;
 RT "A chorionic gonadotropin-sensitive mutation in the
 RT follicle-stimulating hormone receptor as a cause of familial
 RT gestational spontaneous ovarian hyperstimulation syndrome.";
 RL New Engl. J. Med. 349:753-759(2003).
 RN [12]
 RP VARIANT OHSS ASN-567.
 RX MEDLINE=22812037; PubMed=12930928;
 RA Smits G., Olatundusun O., Delbaere A., Pierson R., Vassart G.,
 RA Costagliola S.;
 RT "Ovarian hyperstimulation syndrome due to a mutation in the
 RT follicle-stimulating hormone receptor.";
 RL New Engl. J. Med. 349:760-766(2003).
 CC -1- FUNCTION: Receptor for follicle stimulating hormone. The activity
 CC of this receptor is mediated by G proteins which activate
 CC adenylyate cyclase.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Long;
 CC IsoId=P23945-1; Sequence=Displayed;
 CC Name=Short;
 CC IsoId=P23945-2; Sequence=VSP_001953, VSP_001954;
 CC -1- TISSUE SPECIFICITY: Sertoli cells and ovarian granulosa cells.
 CC -1- DISEASE: Defects in FSHR are a cause of ovarian hyperstimulation
 CC syndrome (OHSS) (MIM:608115). OHSS is a disorder which occurs
 CC either spontaneously or most often as an iatrogenic complication
 CC of ovarian stimulation treatments for in vitro fertilization. The
 CC clinical manifestations vary from abdominal distention and
 CC discomfort to potentially life-threatening, massive ovarian
 CC enlargement and capillary leak with fluid sequestration.
 CC Pathologic features of this syndrome include the presence of
 CC multiple serous and hemorrhagic follicular cysts lined by
 CC luteinized cells, a condition called hyperreactio luteinalis.
 CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC FSH/LSH subfamily.
 CC -1- SIMILARITY: Contains 5 leucine-rich (LRR) repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M65085; AAA52477.1; -;
 CC EMBL; S59900; AAB26480.1; -;
 CC EMBL; M95489; AAA52478.1; -;
 CC EMBL; X68044; CAA48179.1; -;
 CC EMBL; S73199; AAB32071.1; -;
 CC EMBL; S73526; AAB32225.1; -;
 CC PIR; I57661; ORHUT.
 CC PDB; 1XUN; 15-MAY-97.
 CC Genew; HGNC:3969; FSHR.
 CC MIM; 136435; -;
 CC MIM; 608115; -;
 CC GO; GO:0016021; C:integral to membrane; TAS.
 CC GO; GO:0005886; C:plasma membrane; TAS.
 CC GO; GO:0004963; F:follicle stimulating hormone receptor activity; TAS.
 CC GO; GO:0007292; P:female gamete generation; TAS.
 CC GO; GO:0008585; P:female gonad development; TAS.
 CC GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; TAS.

DR GO; GO:0007283; P:spermatogenesis; TAS.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR; 3.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR SMART; SM00013; LRRNT; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing;
KW 3D-structure; Polymorphism; Disease mutation.
FT SIGNAL 1 17
FT CHAIN 18 695
FT DOMAIN 18 366
FT TRANSMEM 367 387
FT DOMAIN 388 398
FT TRANSMEM 399 421
FT DOMAIN 422 443
FT TRANSMEM 444 465
FT DOMAIN 466 485
FT TRANSMEM 486 508
FT DOMAIN 509 528
FT TRANSMEM 529 550
FT DOMAIN 551 573
FT TRANSMEM 574 597
FT DOMAIN 598 608
FT TRANSMEM 609 630
FT DOMAIN 631 695
FT REPEAT 69 93
FT REPEAT 119 143
FT REPEAT 170 192
FT REPEAT 193 216
FT REPEAT 218 240
FT DISULFID 442 517
FT CARBOHYD 191 191
FT CARBOHYD 199 199
FT CARBOHYD 293 293
FT CARBOHYD 318 318
FT VARSPLIC 224 285
BY SIMILARITY.
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
Missing (in isoform short).
/FTid=VSP_001953.

Query Match 4.6%; Score 107.5; DB 1; Length 695;
Best Local Similarity 17.0%; Pred. No. 1.1;
Matches 84; Conservative 77; Mismatches 148; Indels 185; Gaps 17;

QY 79 HEDALTNTNIIIFVAIHREHYTSLMDRLHLVG-----KILIDVSNMR 121
DB 98 HEIRIEKANNL-LYINPEAFQNLNLYLLISNTGIKHLPDVHKIHSLOKYLDDIDNIN 156
QY 122 INQYPSNAEYLAFLPDSLIVKGFNVSAWALQBPKASRQVYICSNNIQARQVIEL 181
DB 157 IH-----TIERNSFVGLSFESVILWL-----NKNIGIQEIHNC----- 189
QY 182 ARQINFIPIDLGLSSAREIENLPLRLFTLWRGPVVAISLATFFLYSFYRDVIHPYAR 241
DB 190 ---FNGTQDLDELNLSDNNNLDELNDVFHGASGFVLDISRTRIHSLPSYGLENLKLRA 246
QY 242 NQGSDFYKIPRIEIVNKTLP-PIVAITLLSLVY-----PWLETWLQCRKQL 305
DB 247 RSTYNLKLPL-----TLEKVALMEASLTYPSHCCAFANWRQISELHPICNKSILROE 300
QY 272 -----LAGLLAAAYQLYGTKYRF-----PWLETWLQCRKQL 305
DB 301 VDYMTQTRGQRSSLAEDNESSYSGPDMTYTEFDYDLCEVVDVTCSPKPDANPCEDIM 360
QY 306 G-----LISFEFAMVHA-----YSLCLPMRSEERYFLNMAYQOVHANTEN 347
DB 361 GYNILRLVLIWFISILAITGNITIVLITTSQYKLTVP-----RFLMCNLAFAD----- 408
QY 348 SWNEEEVVRLEMYISFGIMSLGLSLAVTSIPSVSNALNWRB-----FSFIQSTL 398

DB 409 -----LCIGIYLLLIASVDIHTKSQYHNVAIDWQTGAGCDAGFTVFASSEL 455
QY 399 G---YVALLISTFH-----VLIGWKRAFEERYRFTYTPNFV-LA 435
DB 456 SVYTLTAITLERWHTITHAMQDCKYQLRHAASVMWGMWIFAFALFPFIGISSYMKVS 515
QY 436 LVLPSTVILDLIQL 449
DB 516 ICLPMDIDSPISQL 529

RESULT 10

FSHR_PIG STANDARD; PRT; 695 AA.
ID FSHR_PIG
AC P49059; 077514;
DT 01-FEB-1996 (Rel. 33, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Follicle stimulating hormone receptor precursor (FSH-R) (Follictrophin receptor).
GN FSHR.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=96011644; PubMed=7590277;
RA Remy J.J., Lahbib-Mansais Y., Yerle M., Bozon V., Couture L., Pajot E., Grebert D., Salesse R.;
RT "The porcine follictrophin receptor: cDNA cloning, functional expression and chromosomal localization of the gene."
RL Gene 163:257-261 (1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Wang Y.F., Meyer K.B., Schmidt K., Wan S.J., Degen S.J.F., la Barbera A.R.;
RT "Porcine follicle-stimulating hormone receptor."
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Receptor for follicle stimulating hormone. The activity of this receptor is mediated by G proteins which activate adenylylate cyclase.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC -1- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; L31966; AAA86933.1; -
DR EMBL; AF025377; AAC24981.1; -
DR HSSP; P23945; 1XUN.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR; 4.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR SMART; SM00013; LRRNT; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal; Phosphorylation; Repeat; Leucine-rich repeat.


```
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 695 FOLLICLE STIMULATING HORMONE RECEPTOR.
FT DOMAIN 18 366 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 367 387 1 (POTENTIAL).
FT DOMAIN 388 398 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 399 421 2 (POTENTIAL).
FT DOMAIN 422 443 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 444 465 3 (POTENTIAL).
FT DOMAIN 466 485 4 (POTENTIAL).
FT TRANSMEM 509 528 5 (POTENTIAL).
FT DOMAIN 529 550 5 (POTENTIAL).
FT TRANSMEM 551 573 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 574 597 6 (POTENTIAL).
FT TRANSMEM 598 608 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 609 630 7 (POTENTIAL).
FT DOMAIN 631 695 CYTOPLASMIC (POTENTIAL).
FT REPEAT 44 68 LRR 1.
FT REPEAT 69 93 LRR 2.
FT REPEAT 119 143 LRR 3.
FT REPEAT 170 192 LRR 4.
FT REPEAT 193 216 LRR 5.
FT REPEAT 218 240 LRR 6.
FT DISULFID 442 517 BY SIMILARITY.
FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 2 2 S -> A (IN REF. 1).
FT CONFLICT 13 13 T -> S (IN REF. 1).
FT CONFLICT 60 60 V -> A (IN REF. 1).
FT CONFLICT 166 166 V -> M (IN REF. 1).
FT CONFLICT 215 215 Q -> H (IN REF. 1).
FT CONFLICT 247 247 K -> R (IN REF. 1).
FT CONFLICT 257 257 S -> T (IN REF. 1).
FT CONFLICT 334 334 D -> N (IN REF. 1).
FT CONFLICT 349 349 E -> K (IN REF. 1).
FT CONFLICT 352 352 T -> A (IN REF. 1).
FT CONFLICT 383 383 V -> E (IN REF. 1).
FT CONFLICT 407 407 A -> T (IN REF. 1).
FT CONFLICT 421 421 V -> I (IN REF. 1).
FT CONFLICT 427 427 T -> S (IN REF. 1).
FT CONFLICT 435 435 D -> N (IN REF. 1).
FT CONFLICT 483 483 L -> V (IN REF. 1).
FT CONFLICT 550 550 T -> I (IN REF. 1).
FT CONFLICT 586 586 A -> V (IN REF. 1).
FT CONFLICT 607 607 S -> L (IN REF. 1).
FT CONFLICT 691 691 R -> H (IN REF. 1).
SQ SEQUENCE 695 AA; 78172 MW; E9EBEDB29C79C450 CRC64;

Query Match 4.4%; Score 104.5; DB 1; Length 695;
Best Local Similarity 16.5%; Pred. No. 1.9;
Matches 91; Conservative 91; Mismatches 169; Indels 199; Gaps 21;
```

```
QY 268 -----SLVYLAGLAAAYQLYGTKRRF-----PWLFTWLQCRKOLG--- 306
DB 305 TQARGQRVSLAEDGESSLAKEFDTMSEFYDLCEVVDVICSPEPTENPCEDIMGDI 364
QY 307 -LLSFFFAVHVA-----YSLCPMRRSERYLFLLMAYQQVHANIENSWNE 351
DB 365 LRVLIWFISILAITGNIIVLVILTSQYKLTVP-----RFLMCNLAFAD----- 408
QY 352 EEVWRIEMYISFGIMSLGLSLAVTSIPSVSNALNRE-----SFISQTLG--- 399
DB 409 -----LCIGIYLLIASVDIHTKTQYHNYAIDWQTGAGCDAGFFTVFASLSVYT 459
QY 400 YVALLISTFH-----VLTYGWRKAFEEERYRFTPPNEV-LALVLP 439
DB 460 LTAITLERWHTITHAMQLOCKVQLRHAASIMLVGWTFAFTVALPFIIGISSYMKVSIQLP 519
QY 440 SIVLIDLLQL 449
DB 520 MDIDSPLSQL 529
```

```
RESULT 11
SOTB_EC057 STANDARD; PRT; 396 AA.
AC P58529;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Sugar efflux transporter.
GN SOTB OR 22173 OR ECS2135.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoudis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Toke T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -1- FUNCTION: Involved in the efflux of sugars. The physiological role
CC may be the reduction of the intracellular concentration of toxic
CC sugars or sugar metabolites. Transports L-arabinose and to a
CC lesser extent IPTG. Seems to contribute to the control of the
CC arabinose regulon (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Probable).
CC -1- SIMILARITY: Belongs to major facilitator superfamily. SotB
CC (TC 2.A.1.2) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
```

```

CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL; AE005352; AAC56236.1; -.
DR EMBL; AP002557; BAB35558.1; -.
DR PIR; G90895; G90895.
DR PIR; H85721; H85721.
DR HAMAP; MF_00517; -.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR PROSITE; PS00850; MFS; 1.
KW Transport; Sugar transport; Transmembrane; Inner membrane;
KW Complete proteome.
FT TRANSMEM 15 35 POTENTIAL.
FT TRANSMEM 50 70 POTENTIAL.
FT TRANSMEM 81 101 POTENTIAL.
FT TRANSMEM 103 123 POTENTIAL.
FT TRANSMEM 136 156 POTENTIAL.
FT TRANSMEM 170 190 POTENTIAL.
FT TRANSMEM 209 229 POTENTIAL.
FT TRANSMEM 246 266 POTENTIAL.
FT TRANSMEM 275 295 POTENTIAL.
FT TRANSMEM 299 319 POTENTIAL.
FT TRANSMEM 333 353 POTENTIAL.
FT TRANSMEM 364 384 POTENTIAL.
SQ SEQUENCE 396 AA; 42568 MW; CB6A209B0BE6D4F0 CRC64;

Query Match      4.4%; Score 104; DB 1; Length 396;
Best Local Similarity 18.8%; Pred. No. 1;
Matches 85; Conservative 72; Mismatches 148; Indels 146; Gaps 20;

QY 31 VTWGVIGSGDFAKSLTIRLCRGYHNVIGSRNPKFSEFFPHVVDVTHHEDALTNTNIF 90
Db 31 VPVGLL--SDIAQSFHMQTAQVGIMLTIVAVVALMSLPFMLMTSQVERKLLICLFVVF 88
QY 91 VAIHREHYTSLWDLRHLLVGKILIDVSNMNRINOYPESNAEYLASLPDLSLIVGFNVVS 150
Db 89 IASHVLSFLS-WSFTVLVISRI-----GVAFAPHAIF-----WSITA 123
QY 151 AMALQLGP--KDSARQYVICSNNIQARQVIELAR-----QLNFIPIDLSL----- 195
Db 124 SLAIRMAPAGKRAQALSLATGTALAMVLGLGRIVGQYFGWRMTFFAIGICALITLLC 183
QY 196 -----SSAREIENLPLRLFTLWRGPVVAISLAT-----FFFLYSFVRDVIHPY 239
Db 184 LIKLLPLLPSEHSGSLKSLPL---LFRPALMSITLLTVVVTATHTATYSY----IEPF 235
QY 240 ARN---QOSDFYKIPFIEIVNKTLPVIAITLLSVYLAGLLAAAYQLYGTYRFRFPWLE 296
Db 236 VQNIAGFSANF-----ATALLLLGGAGIIGSVIFGKLGNYAS----- 274
QY 297 TWLQCRKQLGLLSFFFAVAVAYSLCLPMRRSERYLFLMAYQQOVHANIEGNNEEVWR 356
Db 275 -----ALVSTAILLVCLALLLPANSE-----IHGVLSTF-----WG 309
QY 357 IEMVYSFGIMSLGLSL-----LAVTSIPSVSN-----ALNMRFSFIQST 397
Db 310 IANMTIGLMQVKVLTALAPDADVAMALLFSGIFNIGIGAGALVGNVSLHMS-----MSM 364
QY 398 LGYVALLISTFHVLIYG-----WKRAPBEE 422
Db 365 IGYVG-TVPAFAPALIMSTIFFRWPTLLEQ 394

RESULT 12
SOTB ECOLI STANDARD; PRT; 396 AA.
AC P31122; P76883; P77353;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Sugar efflux transporter.
GN SOTB OR B1528.

```

```

OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
RA Sempel G., Seki Y., Sivasubram S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.,
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 28.0-40.1 min region on the linkage map."
RL DNA Res. 3:363-377(1996).
RN [3]
RP SEQUENCE OF 217-396 FROM N.A.
RX MEDLINE=93186717; PubMed=8383113;
RA Cohen S.P., Haechler H., Levy S.B.
RT "Genetic and functional analysis of the multiple antibiotic
resistance (mar) locus in Escherichia coli."
RL J. Bacteriol. 175:1484-1492(1993).
RN [4]
RP CHARACTERIZATION.
RC STRAIN=SB0;
RX MEDLINE=99194728; PubMed=10094697;
RA Bost S., Silva F., Belin D.
RT "Transcriptional activation of ydeA, which encodes a member of the
major facilitator superfamily, interferes with arabinose accumulation
and induction of the Escherichia coli arabinose PBAD promoter."
RL J. Bacteriol. 181:2185-2191(1999).
RN [5]
RP CHARACTERIZATION.
RC STRAIN=JS219;
RX MEDLINE=99369894; PubMed=10436792;
RA Carole S., Pichoff S., Bouche J.-P.
RT "Escherichia coli gene ydeA encodes a major facilitator pump which
exports L-arabinose and isopropyl-beta-D-thiogalactopyranoside."
RL J. Bacteriol. 181:5123-5125(1999).
CC -1- FUNCTION: Involved in the efflux of sugars. The physiological role
may be the reduction of the intracellular concentration of toxic
sugars or sugar metabolites. Transports L-arabinose and to a
lesser extent IPTG. Seems to contribute to the control of the
arabinose regulon.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
(Probable).
CC -1- SIMILARITY: Belongs to major facilitator superfamily. SotB
(TC 2.A.1.2) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.lsb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL; AE000250; AAC74601.1; -.
DR EMBL; D90795; BAA15210.1; -.
DR EMBL; D90796; BAA15218.1; -.
DR EMBL; D90797; BAA15230.1; -.

```


DR EMBL; M96235; -; NOT_ANNOTATED_CDS.
DR PIR; C64907; C64907.
DR EcoGene; EG1636; socB.
DR HAMAP; MF_00517; -; 1.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR Pfam; PF00083; Sugar_tr; 1.
DR PROSITE; PS50850; MFS; 1.
KW Transport; Sugar transport; Transmembrane; Inner membrane;
KW Complete proteome.
FT TRANSMEM 15 35 POTENTIAL.
FT TRANSMEM 50 70 POTENTIAL.
FT TRANSMEM 81 101 POTENTIAL.
FT TRANSMEM 103 123 POTENTIAL.
FT TRANSMEM 136 156 POTENTIAL.
FT TRANSMEM 170 190 POTENTIAL.
FT TRANSMEM 209 229 POTENTIAL.
FT TRANSMEM 246 266 POTENTIAL.
FT TRANSMEM 275 295 POTENTIAL.
FT TRANSMEM 299 319 POTENTIAL.
FT TRANSMEM 333 353 POTENTIAL.
FT TRANSMEM 364 384 POTENTIAL.
SQ SEQUENCE 396 AA; 42538 MW; CB6A34CA4EE6D4F0 CRC64;

Query Match 4.4%; Score 104; DB 1; Length 396;
Best Local Similarity 18.8%; Pred. No. 1;
Matches 85; Conservative 72; Mismatches 148; Indels 146; Gaps 20;

QY 31 VTVGVIGSGDFAKSLTIRLCGYHVVIGSNPKFASEFPFHVVDVTHEDALTKNTIF 90
DB 31 VPEVGLL--SDIAQSFHMQTAQVGMILTIVANVVALMSLPFMLMTSQYERRKLLICLFVVF 88
QY 91 VAHREHYTSLMDLRHLVSKILIDVSNMNRINQYPSNAEYLASLPDLSLVKGFNVVS 150
DB 89 IASHVLSFLS-WSFTVLVISRI-----GVAFAHAIF-----WSITA 123
QY 151 AWALQLGP--KQASROYVICSNNIQARQVIELAR-----QLNFIIDLSL----- 195
DB 124 SLAIRMAPAGKRAQALSLIATGTALAMVLGLPLGRIVGQYFGWRMTFAIGALITLLC 183
QY 196 -----SSAREIENLPLRLFTLWRGPVVAISLAT-----FFELYSFVRDVIHPY 239
DB 184 LKLLPLLPSEHSGSLKSLPL---LFRRPALMSIYLLTVVVTAAHYTASY---IEPF 235
QY 240 ARN---QOSDFYKPIETIVNKTLPVIAITLSLYLAGLLAAAYQLVYGTXYRFPWLE 296
DB 236 VQNIAGFSANF-----ATALLLLGAGIIGSVIFGKLGNYAS----- 274
QY 297 TWLQCRKQLGLLSFFRANVAVSLCLPMRSEERYLFLNWAYQOVHANIEENSWNEEVR 356
DB 275 -----ALVSTAIALLVCLALLPRAANSE-----IHLGVLSIF-----WG 309
QY 357 IEMYISFGIMSLGLSL-----LAVTSIPSVSN-----ALNWRREFSFIQST 397
DB 310 IAMMILGLGMQVKVLAIPADTVAMALFSGIFNIGIGAGALVGNQVSLHWS-----MSM 364
QY 398 LGYVALLISTFHVLIYG-----WKRAFEFE 422
DB 365 IGYVG-AVPAFAALIWSIIFRRWPEVTLLEQ 394

RESULT 13
NPT1_MOUSE
ID NPT1_MOUSE STANDARD; PRT; 465 AA.
AC Q61983;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Renal sodium-dependent phosphate transport protein 1 (sodium/phosphate cotransporter 1) (Na(+)/Pi cotransporter 1) (Renal sodium-phosphate transporter protein 1) (Renal Na(+)-dependent phosphate cotransporter 1).
GN SLC17A1 OR NPT1.

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=95335846; PubMed=7611445;
RA Chong S.S., Kozak C.A., Liu L., Kristjansson K., Dunn S.T.,
RA Bourdeau J.E., Hughes M.R.;
"Cloning, genetic mapping, and expression analysis of a mouse renal sodium-dependent phosphate cotransporter."
Am. J. Physiol. 268:F1038-F1045(1995).
CC -!- FUNCTION: Important for the reorption of phosphate by the kidney.
CC May be involved in actively transporting phosphate into cells via Na(+)-cotransport in the renal brush border membrane.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Kidney.
CC -!- SIMILARITY: Belongs to the sodium/phosphate cotransporter family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X77241; CA54459.1; -.
DR PIR; S69915; S69915.
DR MGD; MGI:103209; SLC17a1.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR004745; P4_cotranspt.
DR TIGRFAMs; TIGR00894; 2A0114euk; 1.
DR PROSITE; PS50850; MFS; 1.
KW Transport; Symport; Sodium transport; Transmembrane; Glycoprotein.
FT TRANSMEM 79 99 POTENTIAL.
FT TRANSMEM 117 137 POTENTIAL.
FT TRANSMEM 176 196 POTENTIAL.
FT TRANSMEM 199 219 POTENTIAL.
FT TRANSMEM 260 280 POTENTIAL.
FT TRANSMEM 304 324 POTENTIAL.
FT TRANSMEM 337 356 POTENTIAL.
FT TRANSMEM 363 383 POTENTIAL.
FT TRANSMEM 399 419 POTENTIAL.
FT TRANSMEM 429 449 POTENTIAL.
FT TRANSMEM 47 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 47 47 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 56 56 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 465 AA; 51589 MW; C67EE25A2C291EEF CRC64;

Query Match 4.4%; Score 103; DB 1; Length 465;
Best Local Similarity 18.7%; Pred. No. 1.5;
Matches 76; Conservative 64; Mismatches 142; Indels 124; Gaps 16;

QY 110 GKILIDVSNMNRINQYPSNAEYLASLPDLSLVKGFNVSAWALQLGPKDA---SRQVY 166
DB 79 GLISSVFEGMVVQAP---VGYLSGYPMKRITGSSFLSSIMSLIIPAAQVGAALVI 135
QY 167 ICSNNIQARQVIELARQLNFI---PIDLSISAREIENLPLRLFTLWRGPPVV--- 218
DB 136 VCRVLQGIAGTSTGHEIWKWAPLERGLTS-----MTLSGFVM--GPFIVLVS 187
QY 219 -----AISLATFFFLYSFVRDVIHPYARNQOSDF----- 247
DB 188 GFICDLGMPMVFYIFGIVGCVLSLWFFLFDPDPO--HPYMSSEKDYIISLMQAS 245
QY 248 ---YKPIETIVNKTLPVIAITLS----- 268
DB 246 SGRQSLPIKAMKSLPLWAILILNSFAFIWSNSLVTYPTFISTVLHVNVRENGLISSLP 305
QY 269 --LYVLAGLLAAAYQLVYGTXYRFRFPWLETWLQCRKQLGLLSFFRANVAVSLCLPMR 326

Query Match 4.3%; Score 102; DB 1; Length 220;
Best Local Similarity 23.8%; Pred. No. 0.72;
Matches 43; Conservative 25; Mismatches 51; Indels 62; Gaps 7;

QY	267	LSLVYLAGLLAAAYQLYYGT-----KYRFPWLETW-----	298
DB	20	LWLLYTAGFVPAVWTFYLGATGQADPVKTFEHLGLWALRFLITLLVTPMRDLTGIT	79
QY	299	-LQCRKQGLLSFFFAMVAVAYSLCPMRSEKRYLFNNMAYQOVHANIEHSWNNEEYVRI	357
DB	80	LIRYRRALGLLAFYYALMHT-----TYWVLDQGLNLSAITDIVER-	120
QY	358	EMVTSFGTMSLGLLSLLAVTSIPSVSNALNW-----REFSFIQSTLGYVALLISTFHVJ	411
DB	121	RPFTTIGMISLALLVPLALTS-----NNWSIRKLGRWSSLHK-LVYIAIAGSAVHFL	172
QY	412	I 412	
DB	173	M 173	

Search completed: April 16, 2004, 12:39:41
Job time : 20 secs

AC2560
hypothetical protein alr8074 [imported] - Nostoc sp. (strain PCC 7120) plasmid PCC7120ga
C/Species: Nostoc sp. PCC 7120
A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C/Accession: AC2560
R/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, S.; Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
A/Reference number: AB1807; MUID:21595285; PMID:11759840
A/Accession: AC2560
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-211 <KUR>
A/Cross-references: GB:AP003603; PIDN:BA877404.1; PID:g17134847; GSPDB:GN00182
A/Experimental source: strain PCC 7120
C/Genetics:
A/Gene: alr8074
A/Genome: plasmid

Query Match 6.8%; Score 159.5; DB 2; Length 211;
Best Local Similarity 24.9%; Pred. No. 2.4e-05;
Matches 51; Conservative 43; Mismatches 70; Indels 41; Gaps 8;

QY 33 VGVIGSGDFAKSLTIRLCGYHVVIGSRNPKASEFPFHVVDVTHEDALTKTN 87
DB 3 IAFIGIGQVGSALASQLSLDHTVTIARNNSDSVYTAALAKYPELVSSFOEALIAQAE 61
QY 88 IIFVA---IHREHTSLMDLRLHLVGLKILIDVSNMNRIN---QYPESNAEYLASLP 138
DB 62 VIFLATPFTANQALAEVGLDLS---GKILVDCTNPFVGNLTHGLKSEQSGSELVQSFPV 117
QY 139 DSLIVKGFNVSAWALQLGPKDASRQVY-----ICSNNIQARQOVIELARQLNF 187
DB 118 HAKYVKAFTI-----YGFENFEHNTYPPGYNLKPAMLIAGNDVPAKQVSTLCQQLGW 170
QY 188 IPIDLGSLSSAREIENLPLRLFTLW 212
DB 171 EAVDVGNLSMSLHLEHMTL---LW 191

RESULT 3

T50571
probable oxidoreductase [imported] - Streptomyces coelicolor
C/Species: Streptomyces coelicolor
C/Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 28-Jul-2000
C/Accession: T50571
R/Reichenbach, M.; Kieser, H.M.; Denapate, D.; Eichner, A.; Cullum, J.; Kinashi, H.; Hopwood, J.C.
Mol. Microbiol. 21, 77-96, 1996
A/Title: A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb S. coelicolor chromosome
A/Reference number: Z20556; MUID:97000351; PMID:8843436
A/Accession: T50571
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-239 <RED>
A/Cross-references: EMBL:AL133220; PIDN:CAB61708.1
A/Experimental source: strain A3(2)
C/Genetics:
A/Note: SCC75A.08c
C/Superfamily: conserved hypothetical protein MJ1501

Query Match 6.7%; Score 156.5; DB 2; Length 239;
Best Local Similarity 27.9%; Pred. No. 5e-05;
Matches 61; Conservative 41; Mismatches 72; Indels 45; Gaps 10;

QY 25 IKDARKVTGVI-GSGDFAKSLTIRLCGYHVVIGSRNPKASEFPFHVVDVTHHE 80
DB 23 LPDVSGLVGVGLGIGPOGKGLAYRLAKAGQKIVGSRRAERLALAEIGHGVEGADNA 82
QY 81 DALTKNIIFVAIHRE-HYTSLMDLRLHLVGLKILIDVSNMNRINQ-----YPE-SNAE 131
DB 83 ETARRSDVIVAVPMDGHGKTLLESIRAEISGLKLVDCVNPPLGFDKKGAYALKPEBSAAE 142

QY 132 YLASLFPDSLIVKGFNVSAWALQ-----LGPKDASRQVYICSNNIQARQOV 178
DB 143 QAAALLPDSRVAAAFHLSAVLLQDPEIDEIDTDMVVLGEERADVEI-----VQA----- 192
QY 179 IELARQLNFIP-----IDLGLSSAREIENLPLRLFTLWR 213
DB 193 --LAGR---IPGMRGVFAGRLRNAHQVSLVANLISVNR 226

RESULT 4

H69400
conserved hypothetical protein AF1209 - Archaeoglobus fulgidus
C/Species: Archaeoglobus fulgidus
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 29-Sep-1999
C/Accession: H69400
R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, R.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A/Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.F.; Sykes, S.; Smith, H.O.; Woese, C.R.; Venter, J.C.
A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon Pyrococcus furiosus
A/Reference number: A69250; MUID:98049343; PMID:9389475
A/Accession: H69400
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-213 <KLE>
A/Cross-references: GB:AE001021; GB:AE000782; NID:G2689344; PIDN:AAB90038.1; PID:G364937
C/Superfamily: conserved hypothetical protein MJ1501

Query Match 6.3%; Score 147; DB 2; Length 213;
Best Local Similarity 27.5%; Pred. No. 0.00024;
Matches 56; Conservative 44; Mismatches 68; Indels 36; Gaps 11;

QY 37 GSGDFAKSLTIRLCGYHVVIGSRNPKASEFPFHVVDVTHEDALTKTN 88
DB 8 GTGNLGGELALRWGLGYEIVIGSRKLEKAELASDYLVKVGDAITGMREDAETCD- 66
QY 89 IFVAIHREHTSLMD-----LRHLVGLKILID-----VSNMNRINQPE-SNAE 131
DB 67 --VAV---FTIPWEFAFDTAEMLKRLAQKGVISPLVPMKKVGDNF-VYVPEEGSAAE 119
QY 132 YLASLFPDSLIVKGFNVSAWALQLGPKDASRQVYICSNNIQARQOVIELARQLNF-Pi 190
DB 120 KLASVLEESSVVAAYHSIPARRFANLGEFEFWDVPIGDS-GAKEVVVDLTETKISGLRAL 178
QY 191 DLGSLSSAREIENL-PLRLFTLWR 213
DB 179 DAGLSNHLVESLTPLILNVMKR 202

RESULT 5

D69361
conserved hypothetical protein AF0892 - Archaeoglobus fulgidus
C/Species: Archaeoglobus fulgidus
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 29-Sep-1999
C/Accession: D69361
R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, R.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A/Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.F.; Sykes, S.; Smith, H.O.; Woese, C.R.; Venter, J.C.
A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon Pyrococcus furiosus
A/Reference number: A69250; MUID:98049343; PMID:9389475
A/Accession: D69361
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-212 <KLE>
A/Cross-references: GB:AE001042; GB:AE000782; NID:G2689365; PIDN:AAB90348.1; PID:G26497
C/Superfamily: conserved hypothetical protein MJ1501

conserved hypothetical protein MTH248 - Methanobacterium thermoautotrophicum (strain Delta H)
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence #revision 05-Dec-1997 #text_change 29-Sep-1999
C:Accession: A69131
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qi, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.; Kl., S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: function
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: A69131
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-232 <MTH>
A:Cross-references: GB:AE00081; GB:AE000666; NID:g2621287; PIDN:AAB84754.1; PID:g262129
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH248
A:Start codon: TTG
C:Superfamily: conserved hypothetical protein MJ1501

Query Match	5.1%;	Score 120.5;	DB 2;	Length 232;
Best Local Similarity	23.8%;	Pred. No. 0.034;		
Matches 49;	Conservative 42;	Mismatches 90;	Indels 25;	Gaps 6
QY	30	KATVGVIGSGDFAKSLTIRLIRCGYHVIVIGSRNPKFASEFFPHVVDVTHHE-----	80	
Db	8	KIAV-IGGTGDQGLGLALRFAVAGEEVITIGSRDAEKASKASKVLEIAGRDDISVEGATN	66	
QY	81	-DALTKNTIIFVAIH-REHYTSLMDLRHLVLVGKILIDVS-----NNMRINQYPESN	129	
Db	67	PDAASADVVVLTVPLOAQMWTLASIRDQVRDKVLLIDATVPIDSCIGGSAYRYIDLWEGS	126	
QY	130	AEYLASLF---PDSLIIVKGFNVVSAWALQLGPKDASRQVYICSNNIQARQOVTELARQLN	186	
Db	127	AAERAAARFLREQGT RVAAPFNNTISGALLVEVSEPVDDCDLVASDHRDALVAAELAEKID	186	
QY	187	FI-PIDLGSLSSAREIENLPLRLFTL	211	
Db	187	GVRAIECGGLENARIIEKITPLLINL	212	

RESULT 10
 T000121
 hypothetical protein 8 - *Leptospira interrogans*
 C/Species: *Leptospira interrogans*
 C/Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
 C/Accession: T000121
 R/Takahashi, Y.; Akase, K.; Hirano, H.; Fukunaga, M.
 Gene 215, 37-45, 1998
 A/Title: Physical and genetic maps of the *Leptospira interrogans* serovar icterohaemorrhagiae
 A/Reference number: Z14115; MUID:98332717; PMID:9666070
 A/Accession: T000121
 A/Status: preliminary; translated from GB/EMBL/DDBJ
 A/Molecule type: DNA
 A/Residues: 1-216 <TAK>
 A/Cross-references: EMBL:AB010203; NID:g2780763; PIDN:BAA24368.1; PID:g2780765
 A/Experimental source: strain Ictero No.1, Substrain icterohaemorrhagiae

	Query Match	5.1%;	Score 119.5;	DB 2;	Length 216;	
	Best Local Similarity	23.7%;	Pred. No. 0.037;			
	Matches	47;	Conservative	37;	Mismatches	91; Indels 23; Gaps 6;
QY	33 VGVIGSGDPAKSLTIRLCRGHHVVICGRNPKFASEFFPHV--VDVTTHEDALTKTNII	89				
	: :: :: :: :: :: :					
Dd	6 IGIIGSGIVGOTLANGFLKYGAELYKIGTRDPFKLKDMLAKAGASIGSFEEAANFGEIIT	65				
	:: :: :: :: :: :: :					
QY	90 FVAIHREHYTSLMDLRHL-LVGKLIDVSNNMR-----INQYPESNAEYLASTL	136				
	: :: :: :: :: :: :					
Dd	66 VLCKSGSVASEVYLTLSGIDSLNGKTIIDTNPISIEIPPONGVLNFFTSTYNESLMEXLOKQ	125				
	: :: :: :: :: :: :					
QY	137 FPDSLIVKGENVVSAMALQLGP-KDASKROYICSNNTQARQQVIELARQLNFIPIDLGS	194				

```

Db      126  A P K A N E V K C F S S V G S - G L M Y N P O L K G E K P S M F I C G N D S S K Q I K E I L D T F G M D I E D M G K 184
QY      195  L S S A R E I E N L P L R L F T L W 212
Db      185  V E A A R A I E P L C I - - - L W 198

```

RESULT 11
G82642
conserved hypothetical protein XF1737 (imported) - *Xylella fastidiosa* (strain 9a5c)
C/Species: *Xylella fastidiosa*
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C/Accession: G82642
R/anonymous: The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequencing
Nature 406, 151-157, 2000
A/Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.
A/Reference number: A82515; MUID:20365717; PMID:10910347
A/Note: for a complete list of authors see reference number A59328 below
A/Accession: G82642
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-242 <Sim>
A/Cross-references: GB:AE003997; GB:AE003849; NID:g9106805; PIDN:AAF84546.1; GSPDB:GN001
A/Experimental source: strain 9a5c
R/Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; B
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, F
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A/Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitejima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, F
A/Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miraccca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.;
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de sa, R.G.; Santelli, R.V.; Sawasay
M.; Tsuchiko, M.H.; Vallada, H.; Van Sluys, M.A.; Vertovski-Almeida, S.; Vettore, A.L.;
A/Reference number: A59328
A/Contents: annotation
A/Genetics:
A/Gene: XF1737

	Query Match	Best Local Similarity	Score	DB 2;	Length	Matches	Conservative	Mismatches	Indels	Gaps
QY	28	ARKVTVGVIGSGDFAKSLTIRLRGCVHVVIGSRNP----	KFASGFPHVDVTHHEDAL	83						
Db	30	AAPMRIGVIGAGSLGTYRRLWVKAGHEVMFSRRNPKLBAMARELEPR-ASVGQPLAAT	88							
QY	84	TKTNIIFVAIHREHYTSLEW-DLRHLVLGKLIDVSNMNRINQ--YBESN----	AEYLAS	135						
Db	89	EFGTVLLLVLPFEALPYGRDLRSAYRGKIIVLDSTNPWGASSADVYREARELGVAQTIVK	148							
QY	136	LFPDSILVKGFNVSAMWLQLGPKDASRQ-----VYICSNNIQARQVIELARQLNFPI	190							
Db	149	YMPGARLVRAFSAVDATAVE--TSASRRGGRIQMPLASDDAEAMKYAEGLVRDAGCDPV	205							
QY	191	DLGSLSAREIE	202							
Db	206	IYGNLAAASFO	217							

RESULT 12
AB3182
conserved hypothetical protein Atu5183 [imported] - *Agrobacterium tumefaciens* (strain C58)
C/Species: *Agrobacterium tumefaciens*
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C/Accession: AB3182
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, J.;
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayarin, T.; Levy, R.; Li, M.; McClellan,
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 12:40:42 ; Search time 48 Seconds

(without alignment)
2607.409 Million cell updates/sec

Title: US-09-455-486-6

Perfect score: 2351

Sequence: 1 MESISMGSPKSLSEICLPN.....ALVLPSTVILDLLOLCRYPD 454

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2351	100.0	454	10	US-09-888-257A-10
2	2351	100.0	454	12	US-10-455-822-3
3	2351	100.0	454	12	US-10-455-822-7
4	2351	100.0	454	12	US-10-455-822-9
5	2351	100.0	454	12	US-10-455-822-19
6	2351	100.0	454	12	US-10-455-822-21
7	2351	100.0	454	12	US-10-455-822-23
8	2351	100.0	454	12	US-10-455-822-25
9	2351	100.0	454	12	US-10-455-822-31
10	2351	100.0	454	12	US-10-455-822-33
11	2351	100.0	454	12	US-10-455-822-35
12	2351	100.0	454	12	US-10-455-822-37
13	2351	100.0	454	12	US-10-455-822-39
14	2351	100.0	454	12	US-10-455-822-78
15	2351	100.0	454	12	US-10-455-822-121

16	2351	100.0	454	12	US-10-455-822-160	Sequence 160, App
17	2351	100.0	454	12	US-10-455-822-161	Sequence 161, App
18	2351	100.0	454	12	US-10-455-822-162	Sequence 162, App
19	2351	100.0	454	12	US-10-455-822-166	Sequence 166, App
20	2351	100.0	454	12	US-10-455-822-167	Sequence 167, App
21	2351	100.0	454	12	US-10-455-822-168	Sequence 168, App
22	2351	100.0	454	12	US-10-455-822-173	Sequence 173, App
23	2351	100.0	454	12	US-10-455-822-179	Sequence 179, App
24	2351	100.0	454	12	US-10-455-822-185	Sequence 185, App
25	2351	100.0	454	12	US-10-455-822-191	Sequence 191, App
26	2351	100.0	454	12	US-10-455-822-193	Sequence 193, App
27	2351	100.0	454	14	US-10-165-044-8	Sequence 8, Appl
28	2351	100.0	454	15	US-10-239-607-29	Sequence 29, Appl
29	2351	100.0	454	15	US-10-295-027-574	Sequence 574, App
30	2351	100.0	454	15	US-10-392-190-2	Sequence 2, Appl
31	2348	99.9	454	15	US-10-104-047-3155	Sequence 3155, Ap
32	2347	99.8	454	12	US-10-455-822-29	Sequence 29, Appl
33	2340	99.5	454	12	US-10-455-822-27	Sequence 27, Appl
34	2340	99.5	454	12	US-10-455-822-84	Sequence 84, Appl
35	2336	99.4	454	12	US-10-455-822-85	Sequence 85, Appl
36	2294	97.6	444	12	US-10-455-822-88	Sequence 88, Appl
37	2294	97.6	444	12	US-10-455-822-90	Sequence 90, Appl
38	2294	97.6	444	12	US-10-455-822-92	Sequence 92, Appl
39	2294	97.6	490	9	US-09-802-520-1	Sequence 1, Appl
40	2294	97.6	490	12	US-10-455-822-13	Sequence 13, Appl
41	2294	97.6	490	12	US-10-455-822-17	Sequence 17, Appl
42	2294	97.6	490	12	US-10-455-822-51	Sequence 51, Appl
43	2294	97.6	490	12	US-10-455-822-53	Sequence 53, Appl
44	2294	97.6	490	12	US-10-455-822-55	Sequence 55, Appl
45	2294	97.6	490	12	US-10-455-822-57	Sequence 57, Appl

ALIGNMENTS

RESULT 1
US-09-888-257A-10
Sequence 10, Application US/09888257A
Public Information No. US20030060612A1
GENERAL INFORMATION:
APPLICANT: Goddard, Audrey
APPLICANT: Goddard, Paul J.
APPLICANT: Guirney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Polakis, Paul
APPLICANT: Smith, Victoria
APPLICANT: Wood, William I.
APPLICANT: Wu, Thomas D.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
FILE REFERENCE: P5002R1
CURRENT APPLICATION NUMBER: US/09/888, 257A
PRIOR FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: ~~US 60/403, 592~~ NO
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: ~~US 60/403, 592~~ NO
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: ~~US 60/403, 592~~ NO
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: ~~US 60/403, 592~~ NO
PRIOR FILING DATE: 1998-10-08
PRIOR APPLICATION NUMBER: ~~US 60/403, 592~~ NO
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: ~~US 60/403, 592~~ NO
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: ~~US 60/403, 592~~ NO
PRIOR FILING DATE: 1999-09-01
PRIOR APPLICATION NUMBER: ~~US 60/403, 592~~ NO
PRIOR FILING DATE: 2000-02-18
PRIOR APPLICATION NUMBER: ~~US 60/403, 592~~ NO
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: ~~US 60/403, 592~~ NO

Pro 23203
disclosed 4/24/2000
60/235 451

PRIOR FILING DATE: 2000-03-30
 PRIOR APPLICATION NUMBER: PCT/US00/23328
 PRIOR FILING DATE: 2000-08-24
 PRIOR APPLICATION NUMBER: PCT/US00/32678
 PRIOR FILING DATE: 2000-12-01
 PRIOR APPLICATION NUMBER: PCT/US01/06520
 PRIOR FILING DATE: 2001-02-28
 PRIOR APPLICATION NUMBER: PCT/US01/06666
 NUMBER OF SEQ ID NOS: 10
 SEQ ID NO 10
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-09-888-257A-10

Query Match 100.0%; Score 2351; DB 10; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6.1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPNCINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVVIGS 60
 DB 1 MESISMGSPKSLSETCLPNCINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVVIGS 60
 QY 61 RNPKEASEFPFPHVDVTHHEDALTKNTIIFVAIHREHYTSLMDLRHLLVGKILLIDVSNM 120
 DB 61 RNPKEASEFPFPHVDVTHHEDALTKNTIIFVAIHREHYTSLMDLRHLLVGKILLIDVSNM 120
 QY 121 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQGPKDASROYTICSNNIQARQVIE 180
 DB 121 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQGPKDASROYTICSNNIQARQVIE 180
 QY 181 LARQNFIPIDLSLSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
 DB 181 LARQNFIPIDLSLSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
 QY 241 RNOQSDFYKIPIEIVNKTPIVAITLLSLVYLAGLLAAAYQOLYYGTYKRRFPWLETWLQ 300
 DB 241 RNOQSDFYKIPIEIVNKTPIVAITLLSLVYLAGLLAAAYQOLYYGTYKRRFPWLETWLQ 300
 QY 301 CRKQGLLSFFPAMVHVAISLCLPMRSEERYFLNMAVQOVHANINENSNNEEVWRIEM 360
 DB 301 CRKQGLLSFFPAMVHVAISLCLPMRSEERYFLNMAVQOVHANINENSNNEEVWRIEM 360
 QY 361 ISFGINSLGLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYWKRAFE 420
 DB 361 ISFGINSLGLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYWKRAFE 420
 QY 421 EEYRYFTPPNFVLAIVLPSIVILDLQLCRYPD 454
 DB 421 EEYRYFTPPNFVLAIVLPSIVILDLQLCRYPD 454

RESULT 2

US-10-455-822-3
 Sequence 3, Application US/10455822
 Publication No. US20040048798A1
 GENERAL INFORMATION:
 APPLICANT: Agensys, Inc.
 APPLICANT: Raitano, Arthur B.
 APPLICANT: Ge, Wangmao
 APPLICANT: Jakobovits, Aya
 APPLICANT: Chalilita-Eid, Pia M.
 APPLICANT: Farris, Mary
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer
 FILE REFERENCE: 51158-20016.24
 CURRENT APPLICATION NUMBER: US/10/455,822
 CURRENT FILING DATE: 2003-06-04
 PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
 PRIOR FILING DATE: 2002-12-20
 PRIOR APPLICATION NUMBER: US60/370,387
 PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US60/296,656
 PRIOR FILING DATE: 2001-06-06
 PRIOR APPLICATION NUMBER: US60/091,183
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US60/087,520
 PRIOR FILING DATE: 1998-06-01
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 3
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-455-822-3

Query Match 100.0%; Score 2351; DB 12; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6.1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPNCINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVVIGS 60
 DB 1 MESISMGSPKSLSETCLPNCINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVVIGS 60
 QY 61 RNPKEASEFPFPHVDVTHHEDALTKNTIIFVAIHREHYTSLMDLRHLLVGKILLIDVSNM 120
 DB 61 RNPKEASEFPFPHVDVTHHEDALTKNTIIFVAIHREHYTSLMDLRHLLVGKILLIDVSNM 120
 QY 121 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQGPKDASROYTICSNNIQARQVIE 180
 DB 121 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQGPKDASROYTICSNNIQARQVIE 180
 QY 181 LARQNFIPIDLSLSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
 DB 181 LARQNFIPIDLSLSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
 QY 241 RNOQSDFYKIPIEIVNKTPIVAITLLSLVYLAGLLAAAYQOLYYGTYKRRFPWLETWLQ 300
 DB 241 RNOQSDFYKIPIEIVNKTPIVAITLLSLVYLAGLLAAAYQOLYYGTYKRRFPWLETWLQ 300
 QY 301 CRKQGLLSFFPAMVHVAISLCLPMRSEERYFLNMAVQOVHANINENSNNEEVWRIEM 360
 DB 301 CRKQGLLSFFPAMVHVAISLCLPMRSEERYFLNMAVQOVHANINENSNNEEVWRIEM 360
 QY 361 ISFGINSLGLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYWKRAFE 420
 DB 361 ISFGINSLGLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYWKRAFE 420
 QY 421 EEYRYFTPPNFVLAIVLPSIVILDLQLCRYPD 454
 DB 421 EEYRYFTPPNFVLAIVLPSIVILDLQLCRYPD 454

RESULT 3

US-10-455-822-7
 Sequence 7, Application US/10455822
 Publication No. US20040048798A1
 GENERAL INFORMATION:
 APPLICANT: Agensys, Inc.
 APPLICANT: Raitano, Arthur B.
 APPLICANT: Ge, Wangmao
 APPLICANT: Jakobovits, Aya
 APPLICANT: Chalilita-Eid, Pia M.
 APPLICANT: Farris, Mary
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer
 FILE REFERENCE: 51158-20016.24
 CURRENT APPLICATION NUMBER: US/10/455,822
 CURRENT FILING DATE: 2003-06-04
 PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
 PRIOR FILING DATE: 2002-12-20
 PRIOR APPLICATION NUMBER: US60/370,387
 PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
PRIOR FILING DATE: 2001-09-06
PRIOR APPLICATION NUMBER: US60/296,656
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US60/091,183
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US60/087,520
PRIOR FILING DATE: 1998-06-01
NUMBER OF SEQ ID NOS: 197
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 454
TYPE: PRT
ORGANISM: Homo sapien
US-10-455-822-7

Query Match 100.0%; Score 2351; DB 12; Length 454;
Best Local Similarity 100.0%; Pred. No. 6.1e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISWMSGSPKSLSETCLPNGINGIKDARKVTGVIGSGDPAKSLTIRLRGCHVYIGS 60
DB 1 MESISWMSGSPKSLSETCLPNGINGIKDARKVTGVIGSGDPAKSLTIRLRGCHVYIGS 60
QY 61 RNPKFASEFFPHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGILIDVSNM 120
DB 61 RNPKFASEFFPHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGILIDVSNM 120
QY 121 RINQYPESNAEYLASLPDLSLVKGNVVSAMALQGPKDSRQVYICSNNIQARQVIE 180
DB 121 RINQYPESNAEYLASLPDLSLVKGNVVSAMALQGPKDSRQVYICSNNIQARQVIE 180
QY 181 LARQNFIPIDIGSLSSAREIENLPLRLFTLMRGPVVVAISLATFFFLYSFVRDVHPYA 240
DB 181 LARQNFIPIDIGSLSSAREIENLPLRLFTLMRGPVVVAISLATFFFLYSFVRDVHPYA 240
QY 241 RNOQSDFYKPIEIVNKTLPPIVATITLSLVYLAGLLAAAYQLYYGTKYRRFPFWLETWLQ 300
DB 241 RNOQSDFYKPIEIVNKTLPPIVATITLSLVYLAGLLAAAYQLYYGTKYRRFPFWLETWLQ 300
QY 301 CRKQGLLSFFPAMVHVAYSLCLPMRSEERYLFLMAYQQVHANINENSWNEEVEWRIEMY 360
DB 301 CRKQGLLSFFPAMVHVAYSLCLPMRSEERYLFLMAYQQVHANINENSWNEEVEWRIEMY 360
QY 361 ISFGIMSLGLSLAIVTSIPSVSNALNMBEFSFIQSTLGYVALLISTPHVLIYGMKRAFE 420
DB 361 ISFGIMSLGLSLAIVTSIPSVSNALNMBEFSFIQSTLGYVALLISTPHVLIYGMKRAFE 420
QY 421 EYYRFTYPPNFVLAIVLPSIVILDLQLCRYPD 454
DB 421 EYYRFTYPPNFVLAIVLPSIVILDLQLCRYPD 454

RESULT 4

US-10-455-822-9
Sequence 9, Application US/10455822
Publication No. US20040048798A1
GENERAL INFORMATION:
APPLICANT: Agensys, Inc.
APPLICANT: Raitano, Arthur B.
APPLICANT: Ge, Wangmao
APPLICANT: Jakobovits, Aya
APPLICANT: Challita-Eid, Pia M.
APPLICANT: Faris, Mary
TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer
FILE REFERENCE: 51158-20016.24
CURRENT APPLICATION NUMBER: US/10/455,822
CURRENT FILING DATE: 2003-06-04
PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
PRIOR FILING DATE: 2002-12-20
PRIOR APPLICATION NUMBER: US60/370,387
PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
PRIOR FILING DATE: 2001-09-06
PRIOR APPLICATION NUMBER: US60/296,656
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US60/091,183
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US60/087,520
PRIOR FILING DATE: 1998-06-01
NUMBER OF SEQ ID NOS: 197
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9
LENGTH: 454
TYPE: PRT
ORGANISM: Homo sapien
US-10-455-822-9

Query Match 100.0%; Score 2351; DB 12; Length 454;
Best Local Similarity 100.0%; Pred. No. 6.1e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISWMSGSPKSLSETCLPNGINGIKDARKVTGVIGSGDPAKSLTIRLRGCHVYIGS 60
DB 1 MESISWMSGSPKSLSETCLPNGINGIKDARKVTGVIGSGDPAKSLTIRLRGCHVYIGS 60
QY 61 RNPKFASEFFPHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGILIDVSNM 120
DB 61 RNPKFASEFFPHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGILIDVSNM 120
QY 121 RINQYPESNAEYLASLPDLSLVKGNVVSAMALQGPKDSRQVYICSNNIQARQVIE 180
DB 121 RINQYPESNAEYLASLPDLSLVKGNVVSAMALQGPKDSRQVYICSNNIQARQVIE 180
QY 181 LARQNFIPIDIGSLSSAREIENLPLRLFTLMRGPVVVAISLATFFFLYSFVRDVHPYA 240
DB 181 LARQNFIPIDIGSLSSAREIENLPLRLFTLMRGPVVVAISLATFFFLYSFVRDVHPYA 240
QY 241 RNOQSDFYKPIEIVNKTLPPIVATITLSLVYLAGLLAAAYQLYYGTKYRRFPFWLETWLQ 300
DB 241 RNOQSDFYKPIEIVNKTLPPIVATITLSLVYLAGLLAAAYQLYYGTKYRRFPFWLETWLQ 300
QY 301 CRKQGLLSFFPAMVHVAYSLCLPMRSEERYLFLMAYQQVHANINENSWNEEVEWRIEMY 360
DB 301 CRKQGLLSFFPAMVHVAYSLCLPMRSEERYLFLMAYQQVHANINENSWNEEVEWRIEMY 360
QY 361 ISFGIMSLGLSLAIVTSIPSVSNALNMBEFSFIQSTLGYVALLISTPHVLIYGMKRAFE 420
DB 361 ISFGIMSLGLSLAIVTSIPSVSNALNMBEFSFIQSTLGYVALLISTPHVLIYGMKRAFE 420
QY 421 EYYRFTYPPNFVLAIVLPSIVILDLQLCRYPD 454
DB 421 EYYRFTYPPNFVLAIVLPSIVILDLQLCRYPD 454

RESULT 5

US-10-455-822-19
Sequence 19, Application US/10455822
Publication No. US20040048798A1
GENERAL INFORMATION:
APPLICANT: Agensys, Inc.
APPLICANT: Raitano, Arthur B.
APPLICANT: Ge, Wangmao
APPLICANT: Jakobovits, Aya
APPLICANT: Challita-Eid, Pia M.
APPLICANT: Faris, Mary
TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer
FILE REFERENCE: 51158-20016.24
CURRENT APPLICATION NUMBER: US/10/455,822
CURRENT FILING DATE: 2003-06-04
PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
PRIOR FILING DATE: 2002-12-20
PRIOR APPLICATION NUMBER: US60/370,387
PRIOR FILING DATE: 2002-04-05


```

; PRIOR APPLICATION NUMBER: US60/317,840
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: US60/296,656
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US60/091,183
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US60/087,520
; PRIOR FILING DATE: 1998-06-01
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-455-822-19
```

```

Query Match      100.0%; Score 2351; DB 12; Length 454;
Best Local Similarity 100.0%; Pred. No. 6.1e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVIGS 60
Db      1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVIGS 60
QY      61 RNPKEASEFFPHVVDVTHHEDALTNTNIIFVAIHREHYTSLMDLRHLVYKILIDVSNM 120
Db      61 RNPKEASEFFPHVVDVTHHEDALTNTNIIFVAIHREHYTSLMDLRHLVYKILIDVSNM 120
QY      121 RINQYPESNAEYLASLPDLSLVKGFNVSAWALQGPKDASRQVYICSNNIQARQVIE 180
Db      121 RINQYPESNAEYLASLPDLSLVKGFNVSAWALQGPKDASRQVYICSNNIQARQVIE 180
QY      181 LARQNFIPIDLSLSAREIENLPRLFTLWGRPVVAISLATFFFLYSFVRDVIHPYA 240
Db      181 LARQNFIPIDLSLSAREIENLPRLFTLWGRPVVAISLATFFFLYSFVRDVIHPYA 240
QY      241 RNOQSDFYKIPIEIVNKTLPVAITLLSLVYLAGLLAAAYQLYGTXYRRPFWLETLQ 300
Db      241 RNOQSDFYKIPIEIVNKTLPVAITLLSLVYLAGLLAAAYQLYGTXYRRPFWLETLQ 300
QY      301 CRKQGLLSFFFAMVHVAISLCLPMRSEERYFLNMAYQOVHANIENSNWEEVWRIEMY 360
Db      301 CRKQGLLSFFFAMVHVAISLCLPMRSEERYFLNMAYQOVHANIENSNWEEVWRIEMY 360
QY      361 ISFGIMSLGLLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYWKRAFE 420
Db      361 ISFGIMSLGLLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYWKRAFE 420
QY      421 EEYRYFTPPNFVLAIVLPSIVILDLQLCRYPD 454
Db      421 EEYRYFTPPNFVLAIVLPSIVILDLQLCRYPD 454
```

```

RESULT 6
US-10-455-822-21
; Sequence 21, Application US/10455822
; Publication No. US20040048798A1
; GENERAL INFORMATION:
; APPLICANT: Agency, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya
; APPLICANT: Chailita-Eid, Pia M.
; APPLICANT: Faris, Mary
; TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer
; FILE REFERENCE: 51158-20016.24
; CURRENT APPLICATION NUMBER: US/10/455,822
; PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US60/370,387
; PRIOR FILING DATE: 2002-04-05
```

```

; PRIOR APPLICATION NUMBER: US60/317,840
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: US60/296,656
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US60/091,183
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US60/087,520
; PRIOR FILING DATE: 1998-06-01
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-455-822-21
```

```

Query Match      100.0%; Score 2351; DB 12; Length 454;
Best Local Similarity 100.0%; Pred. No. 6.1e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVIGS 60
Db      1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVIGS 60
QY      61 RNPKEASEFFPHVVDVTHHEDALTNTNIIFVAIHREHYTSLMDLRHLVYKILIDVSNM 120
Db      61 RNPKEASEFFPHVVDVTHHEDALTNTNIIFVAIHREHYTSLMDLRHLVYKILIDVSNM 120
QY      121 RINQYPESNAEYLASLPDLSLVKGFNVSAWALQGPKDASRQVYICSNNIQARQVIE 180
Db      121 RINQYPESNAEYLASLPDLSLVKGFNVSAWALQGPKDASRQVYICSNNIQARQVIE 180
QY      181 LARQNFIPIDLSLSAREIENLPRLFTLWGRPVVAISLATFFFLYSFVRDVIHPYA 240
Db      181 LARQNFIPIDLSLSAREIENLPRLFTLWGRPVVAISLATFFFLYSFVRDVIHPYA 240
QY      241 RNOQSDFYKIPIEIVNKTLPVAITLLSLVYLAGLLAAAYQLYGTXYRRPFWLETLQ 300
Db      241 RNOQSDFYKIPIEIVNKTLPVAITLLSLVYLAGLLAAAYQLYGTXYRRPFWLETLQ 300
QY      301 CRKQGLLSFFFAMVHVAISLCLPMRSEERYFLNMAYQOVHANIENSNWEEVWRIEMY 360
Db      301 CRKQGLLSFFFAMVHVAISLCLPMRSEERYFLNMAYQOVHANIENSNWEEVWRIEMY 360
QY      361 ISFGIMSLGLLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYWKRAFE 420
Db      361 ISFGIMSLGLLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYWKRAFE 420
QY      421 EEYRYFTPPNFVLAIVLPSIVILDLQLCRYPD 454
Db      421 EEYRYFTPPNFVLAIVLPSIVILDLQLCRYPD 454
```

```

RESULT 7
US-10-455-822-23
; Sequence 23, Application US/10455822
; Publication No. US20040048798A1
; GENERAL INFORMATION:
; APPLICANT: Agency, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya
; APPLICANT: Chailita-Eid, Pia M.
; APPLICANT: Faris, Mary
; TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer
; FILE REFERENCE: 51158-20016.24
; CURRENT APPLICATION NUMBER: US/10/455,822
; PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US60/370,387
; PRIOR FILING DATE: 2002-04-05
```

PRIOR APPLICATION NUMBER: US60/317,840
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US60/296,656
 PRIOR FILING DATE: 2001-06-06
 PRIOR APPLICATION NUMBER: US60/091,183
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US60/087,520
 PRIOR FILING DATE: 1998-06-01
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 23
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-455-822-23

Query Match 100.0%; Score 2351; DB 12; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6.1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHVG 60
 DB 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHVG 60
 QY 61 RNPKFASEFPFPHVDVTHEDALTKNTIFVAIHREHYTSLMDLRHLVGLKILIDVSNM 120
 DB 61 RNPKFASEFPFPHVDVTHEDALTKNTIFVAIHREHYTSLMDLRHLVGLKILIDVSNM 120
 QY 121 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQGPKDASQVYICSNNIQARQVIE 180
 DB 121 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQGPKDASQVYICSNNIQARQVIE 180
 QY 181 LARQNFIPIDLSLSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
 DB 181 LARQNFIPIDLSLSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
 QY 241 RNQSDFYKIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYGTGRRFPFWLETWLQ 300
 DB 241 RNQSDFYKIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYGTGRRFPFWLETWLQ 300
 QY 301 CRKQGLLSFFPAMVHVAVSLCLPMRSEERYFLNMAVYQOVHANIENSNWNEEVRRIEMY 360
 DB 301 CRKQGLLSFFPAMVHVAVSLCLPMRSEERYFLNMAVYQOVHANIENSNWNEEVRRIEMY 360
 QY 361 ISFGIMSLGLSLAATVTSIPSVSNALNWRFSFIQSTLGYVALLISTFVLIYWKRAFE 420
 DB 361 ISFGIMSLGLSLAATVTSIPSVSNALNWRFSFIQSTLGYVALLISTFVLIYWKRAFE 420
 QY 421 BEYYRFYTPPNFVLAIVLPSIVILDLLQLCRYPD 454
 DB 421 BEYYRFYTPPNFVLAIVLPSIVILDLLQLCRYPD 454

RESULT 8
 US-10-455-822-25
 Sequence 25, Application US/10455822
 Publication No. US20040048798A1
 GENERAL INFORMATION:
 APPLICANT: Agensys, Inc.
 APPLICANT: Raitano, Arthur B.
 APPLICANT: Ge, Wangmao
 APPLICANT: Jakobovits, Aya
 APPLICANT: Challita-Bid, Pia M.
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer
 FILE REFERENCE: 51158-20016.24
 CURRENT APPLICATION NUMBER: US/10/455,822
 PRIOR APPLICATION NUMBER: 2003-06-04
 PRIOR FILING DATE: 2002-12-20
 PRIOR APPLICATION NUMBER: US60/370,387
 PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US60/296,656
 PRIOR FILING DATE: 2001-06-06
 PRIOR APPLICATION NUMBER: US60/091,183
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US60/087,520
 PRIOR FILING DATE: 1998-06-01
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 25
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-455-822-25

Query Match 100.0%; Score 2351; DB 12; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6.1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHVG 60
 DB 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHVG 60
 QY 61 RNPKFASEFPFPHVDVTHEDALTKNTIFVAIHREHYTSLMDLRHLVGLKILIDVSNM 120
 DB 61 RNPKFASEFPFPHVDVTHEDALTKNTIFVAIHREHYTSLMDLRHLVGLKILIDVSNM 120
 QY 121 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQGPKDASQVYICSNNIQARQVIE 180
 DB 121 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQGPKDASQVYICSNNIQARQVIE 180
 QY 181 LARQNFIPIDLSLSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
 DB 181 LARQNFIPIDLSLSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
 QY 241 RNQSDFYKIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYGTGRRFPFWLETWLQ 300
 DB 241 RNQSDFYKIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYGTGRRFPFWLETWLQ 300
 QY 301 CRKQGLLSFFPAMVHVAVSLCLPMRSEERYFLNMAVYQOVHANIENSNWNEEVRRIEMY 360
 DB 301 CRKQGLLSFFPAMVHVAVSLCLPMRSEERYFLNMAVYQOVHANIENSNWNEEVRRIEMY 360
 QY 361 ISFGIMSLGLSLAATVTSIPSVSNALNWRFSFIQSTLGYVALLISTFVLIYWKRAFE 420
 DB 361 ISFGIMSLGLSLAATVTSIPSVSNALNWRFSFIQSTLGYVALLISTFVLIYWKRAFE 420
 QY 421 BEYYRFYTPPNFVLAIVLPSIVILDLLQLCRYPD 454
 DB 421 BEYYRFYTPPNFVLAIVLPSIVILDLLQLCRYPD 454

RESULT 9
 US-10-455-822-31
 Sequence 31, Application US/10455822
 Publication No. US20040048798A1
 GENERAL INFORMATION:
 APPLICANT: Agensys, Inc.
 APPLICANT: Raitano, Arthur B.
 APPLICANT: Ge, Wangmao
 APPLICANT: Jakobovits, Aya
 APPLICANT: Challita-Bid, Pia M.
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer
 FILE REFERENCE: 51158-20016.24
 CURRENT APPLICATION NUMBER: US/10/455,822
 PRIOR APPLICATION NUMBER: 2003-06-04
 PRIOR FILING DATE: 2002-12-20
 PRIOR APPLICATION NUMBER: US60/370,387
 PRIOR FILING DATE: 2002-04-05


```

; PRIOR APPLICATION NUMBER: US60/317,840
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: US60/296,656
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US60/091,183
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US60/087,520
; PRIOR FILING DATE: 1998-06-01
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Homo sapien
; US-10-455-822-31

Query Match      100.0%; Score 2351; DB 12; Length 454;
Best Local Similarity 100.0%; Pred. No. 6.1e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVIGS 60
DB      1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVIGS 60

QY      61 RNPKFASEFFPHVVDVTHHEDALTNTNIIFVAIHREHYTSLMDLRHLVSKILIDVSNM 120
DB      61 RNPKFASEFFPHVVDVTHHEDALTNTNIIFVAIHREHYTSLMDLRHLVSKILIDVSNM 120

QY      121 RINQYPESNAEYLASLPDLSLVKGFNVVSAMALQIGPKDASRQVYICSNNIQARQVIE 180
DB      121 RINQYPESNAEYLASLPDLSLVKGFNVVSAMALQIGPKDASRQVYICSNNIQARQVIE 180

QY      181 LARQNFIPIDLSLSAREIENLPLRLFTLWGRGVVVAISLATFFFLYSFVRDVIHPYA 240
DB      181 LARQNFIPIDLSLSAREIENLPLRLFTLWGRGVVVAISLATFFFLYSFVRDVIHPYA 240

QY      241 RNOQSDFYKPIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYYGTXYRRFPPLWLTWQ 300
DB      241 RNOQSDFYKPIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYYGTXYRRFPPLWLTWQ 300

QY      301 CRKQGLLSFFFAMVHVAYSICLPMRSEERYLFNMAYQOVHANIENTSNWNEEVRRIEMY 360
DB      301 CRKQGLLSFFFAMVHVAYSICLPMRSEERYLFNMAYQOVHANIENTSNWNEEVRRIEMY 360

QY      361 ISFGIMSLGLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
DB      361 ISFGIMSLGLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420

QY      421 EBYRYFTPPNFVLAIVLPSIVILDDLQLCRYPD 454
DB      421 EBYRYFTPPNFVLAIVLPSIVILDDLQLCRYPD 454

RESULT 10
US-10-455-822-33
; Sequence 33, Application US/10455822
; Publication No. US20040048798A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Farris, Mary
; TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer
; FILE REFERENCE: 51158-20016.24
; CURRENT APPLICATION NUMBER: US/10/455,822
; PRIOR FILING DATE: 2003-06-04
; PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US60/370,387
; PRIOR FILING DATE: 2002-04-05
```

```

; PRIOR APPLICATION NUMBER: US60/317,840
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: US60/296,656
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US60/091,183
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US60/087,520
; PRIOR FILING DATE: 1998-06-01
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Homo sapien
; US-10-455-822-33

Query Match      100.0%; Score 2351; DB 12; Length 454;
Best Local Similarity 100.0%; Pred. No. 6.1e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVIGS 60
DB      1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVIGS 60

QY      61 RNPKFASEFFPHVVDVTHHEDALTNTNIIFVAIHREHYTSLMDLRHLVSKILIDVSNM 120
DB      61 RNPKFASEFFPHVVDVTHHEDALTNTNIIFVAIHREHYTSLMDLRHLVSKILIDVSNM 120

QY      121 RINQYPESNAEYLASLPDLSLVKGFNVVSAMALQIGPKDASRQVYICSNNIQARQVIE 180
DB      121 RINQYPESNAEYLASLPDLSLVKGFNVVSAMALQIGPKDASRQVYICSNNIQARQVIE 180

QY      181 LARQNFIPIDLSLSAREIENLPLRLFTLWGRGVVVAISLATFFFLYSFVRDVIHPYA 240
DB      181 LARQNFIPIDLSLSAREIENLPLRLFTLWGRGVVVAISLATFFFLYSFVRDVIHPYA 240

QY      241 RNOQSDFYKPIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYYGTXYRRFPPLWLTWQ 300
DB      241 RNOQSDFYKPIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYYGTXYRRFPPLWLTWQ 300

QY      301 CRKQGLLSFFFAMVHVAYSICLPMRSEERYLFNMAYQOVHANIENTSNWNEEVRRIEMY 360
DB      301 CRKQGLLSFFFAMVHVAYSICLPMRSEERYLFNMAYQOVHANIENTSNWNEEVRRIEMY 360

QY      361 ISFGIMSLGLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
DB      361 ISFGIMSLGLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420

QY      421 EBYRYFTPPNFVLAIVLPSIVILDDLQLCRYPD 454
DB      421 EBYRYFTPPNFVLAIVLPSIVILDDLQLCRYPD 454

RESULT 11
US-10-455-822-35
; Sequence 35, Application US/10455822
; Publication No. US20040048798A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Farris, Mary
; TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer
; FILE REFERENCE: 51158-20016.24
; CURRENT APPLICATION NUMBER: US/10/455,822
; PRIOR FILING DATE: 2003-06-04
; PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US60/370,387
; PRIOR FILING DATE: 2002-04-05
```

PRIOR APPLICATION NUMBER: US60/317,840
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US60/296,656
 PRIOR FILING DATE: 2001-06-06
 PRIOR APPLICATION NUMBER: US60/091,183
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US60/087,520
 PRIOR FILING DATE: 1998-06-01
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 35
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-455-822-35

Query Match 100.0%; Score 2351; DB 12; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6.1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCHVIGS 60
 DB 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCHVIGS 60
 QY 61 RNPKEASEFPFHVVDVTHEDALTKNTIIFVAIHREHYTSLMDLRHLVKGKILLIDVSNM 120
 DB 61 RNPKEASEFPFHVVDVTHEDALTKNTIIFVAIHREHYTSLMDLRHLVKGKILLIDVSNM 120
 QY 121 RINQYPSNAEYLAFLPDLSLVKGFNVSAWALQGPKDASROYVIGSNNIQARQVIE 180
 DB 121 RINQYPSNAEYLAFLPDLSLVKGFNVSAWALQGPKDASROYVIGSNNIQARQVIE 180
 QY 181 LARQNFIPIDLSLSAREIENLPLRLFTLMRGPVVAISLATFFLYSFVRDVIHPYA 240
 DB 181 LARQNFIPIDLSLSAREIENLPLRLFTLMRGPVVAISLATFFLYSFVRDVIHPYA 240
 QY 241 RNOQSDFYKIPIEIVNKTLPVIAITLLSLVYLAGLLAAAYQLYGTXYRFPFWLETWLQ 300
 DB 241 RNOQSDFYKIPIEIVNKTLPVIAITLLSLVYLAGLLAAAYQLYGTXYRFPFWLETWLQ 300
 QY 301 CRKQGLLSFFPAMVAVAYSLCLPMRSEERYFLNMAVYQOVHANIENTSNNEEVRRIEM 360
 DB 301 CRKQGLLSFFPAMVAVAYSLCLPMRSEERYFLNMAVYQOVHANIENTSNNEEVRRIEM 360
 QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
 DB 361 ISFGIMSLGLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
 QY 421 EBYRFTYTPNFVLAIVPSIVILLDLQLCRYPD 454
 DB 421 EBYRFTYTPNFVLAIVPSIVILLDLQLCRYPD 454

RESULT 12
 US-10-455-822-37
 Sequence 37, Application US/10455822
 Publication No. US20040048798A1
 GENERAL INFORMATION:
 APPLICANT: Agensys, Inc.
 APPLICANT: Raitano, Arthur B.
 APPLICANT: Ge, Wangmao
 APPLICANT: Jakobovits, Aya
 APPLICANT: Chalilata-Eid, Pia M.
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer
 FILE REFERENCE: 51158-20016.24
 CURRENT APPLICATION NUMBER: US/10/455,822
 PRIOR FILING DATE: 2003-06-04
 PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
 PRIOR FILING DATE: 2002-12-20
 PRIOR APPLICATION NUMBER: US60/370,387
 PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US60/296,656
 PRIOR FILING DATE: 2001-06-06
 PRIOR APPLICATION NUMBER: US60/091,183
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US60/087,520
 PRIOR FILING DATE: 1998-06-01
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 37
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-455-822-37

Query Match 100.0%; Score 2351; DB 12; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6.1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCHVIGS 60
 DB 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCHVIGS 60
 QY 61 RNPKEASEFPFHVVDVTHEDALTKNTIIFVAIHREHYTSLMDLRHLVKGKILLIDVSNM 120
 DB 61 RNPKEASEFPFHVVDVTHEDALTKNTIIFVAIHREHYTSLMDLRHLVKGKILLIDVSNM 120
 QY 121 RINQYPSNAEYLAFLPDLSLVKGFNVSAWALQGPKDASROYVIGSNNIQARQVIE 180
 DB 121 RINQYPSNAEYLAFLPDLSLVKGFNVSAWALQGPKDASROYVIGSNNIQARQVIE 180
 QY 181 LARQNFIPIDLSLSAREIENLPLRLFTLMRGPVVAISLATFFLYSFVRDVIHPYA 240
 DB 181 LARQNFIPIDLSLSAREIENLPLRLFTLMRGPVVAISLATFFLYSFVRDVIHPYA 240
 QY 241 RNOQSDFYKIPIEIVNKTLPVIAITLLSLVYLAGLLAAAYQLYGTXYRFPFWLETWLQ 300
 DB 241 RNOQSDFYKIPIEIVNKTLPVIAITLLSLVYLAGLLAAAYQLYGTXYRFPFWLETWLQ 300
 QY 301 CRKQGLLSFFPAMVAVAYSLCLPMRSEERYFLNMAVYQOVHANIENTSNNEEVRRIEM 360
 DB 301 CRKQGLLSFFPAMVAVAYSLCLPMRSEERYFLNMAVYQOVHANIENTSNNEEVRRIEM 360
 QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
 DB 361 ISFGIMSLGLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
 QY 421 EBYRFTYTPNFVLAIVPSIVILLDLQLCRYPD 454
 DB 421 EBYRFTYTPNFVLAIVPSIVILLDLQLCRYPD 454

RESULT 13
 US-10-455-822-39
 Sequence 39, Application US/10455822
 Publication No. US20040048798A1
 GENERAL INFORMATION:
 APPLICANT: Agensys, Inc.
 APPLICANT: Raitano, Arthur B.
 APPLICANT: Ge, Wangmao
 APPLICANT: Jakobovits, Aya
 APPLICANT: Chalilata-Eid, Pia M.
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer
 FILE REFERENCE: 51158-20016.24
 CURRENT APPLICATION NUMBER: US/10/455,822
 PRIOR FILING DATE: 2003-06-04
 PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
 PRIOR FILING DATE: 2002-12-20
 PRIOR APPLICATION NUMBER: US60/370,387
 PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US60/296,656
 PRIOR FILING DATE: 2001-06-06
 PRIOR APPLICATION NUMBER: US60/091,183
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US60/087,520
 PRIOR FILING DATE: 1998-06-01
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 39
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-455-822-39

Query Match 100.0%; Score 2351; DB 12; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6.1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMWGPSPKSLSETCLPNGINGIKDARKVTGVIIGSGDPAKSLTIRLRGCVHVIIGS 60
 DB 1 MESISMWGPSPKSLSETCLPNGINGIKDARKVTGVIIGSGDPAKSLTIRLRGCVHVIIGS 60
 QY 61 RNPKFASEFFPHVVDVTHHEDALTKNIIIFVAIHREHYTSLMDLRHLVKGILIDVSNM 120
 DB 61 RNPKFASEFFPHVVDVTHHEDALTKNIIIFVAIHREHYTSLMDLRHLVKGILIDVSNM 120
 QY 121 RINQYPSNAEYLASLFPDSLIVKGFNVVSAMALQGPKDASRQVYICSNNIQARQVIE 180
 DB 121 RINQYPSNAEYLASLFPDSLIVKGFNVVSAMALQGPKDASRQVYICSNNIQARQVIE 180
 QY 181 LARQNFIPIDIGSLSSAREIENLPLRLFTLMRGPVVVAISLATFFFLYSFVRDVIHPYA 240
 DB 181 LARQNFIPIDIGSLSSAREIENLPLRLFTLMRGPVVVAISLATFFFLYSFVRDVIHPYA 240
 QY 241 RNQOSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTYRRFPFWLETWLQ 300
 DB 241 RNQOSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTYRRFPFWLETWLQ 300
 QY 301 CRKQGLLSFFFAWVAVASLCLPMRRSERYFLNMAVYQOVHANIENSWNEEVEWRIEMY 360
 DB 301 CRKQGLLSFFFAWVAVASLCLPMRRSERYFLNMAVYQOVHANIENSWNEEVEWRIEMY 360
 QY 361 ISFGIMSLGLSLAIVTSIPSVSNALNWRREFSIQSTLGYVALLISTFHVLIYGMKRAFE 420
 DB 361 ISFGIMSLGLSLAIVTSIPSVSNALNWRREFSIQSTLGYVALLISTFHVLIYGMKRAFE 420
 QY 421 BEYYRFTPPNPFVLAALVPSIVILDLQLCRYPD 454
 DB 421 BEYYRFTPPNPFVLAALVPSIVILDLQLCRYPD 454

RESULT 14
 US-10-455-822-78
 Sequence 78, Application US/10455822
 Publication No. US20040048798A1
 GENERAL INFORMATION:
 APPLICANT: Agensys, Inc.
 APPLICANT: Raitano, Arthur B.
 APPLICANT: Ge, Wangmao
 APPLICANT: Jakobovits, Aya
 APPLICANT: Chalilita-Eid, Pia M.
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer
 FILE REFERENCE: 51158-20016.24
 CURRENT APPLICATION NUMBER: US/10/455,822
 PRIOR FILING DATE: 2003-06-04
 PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
 PRIOR FILING DATE: 2002-12-20
 PRIOR APPLICATION NUMBER: US60/370,387
 PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US60/296,656
 PRIOR FILING DATE: 2001-06-06
 PRIOR APPLICATION NUMBER: US60/091,183
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US60/087,520
 PRIOR FILING DATE: 1998-06-01
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 78
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-455-822-78

Query Match 100.0%; Score 2351; DB 12; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6.1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMWGPSPKSLSETCLPNGINGIKDARKVTGVIIGSGDPAKSLTIRLRGCVHVIIGS 60
 DB 1 MESISMWGPSPKSLSETCLPNGINGIKDARKVTGVIIGSGDPAKSLTIRLRGCVHVIIGS 60
 QY 61 RNPKFASEFFPHVVDVTHHEDALTKNIIIFVAIHREHYTSLMDLRHLVKGILIDVSNM 120
 DB 61 RNPKFASEFFPHVVDVTHHEDALTKNIIIFVAIHREHYTSLMDLRHLVKGILIDVSNM 120
 QY 121 RINQYPSNAEYLASLFPDSLIVKGFNVVSAMALQGPKDASRQVYICSNNIQARQVIE 180
 DB 121 RINQYPSNAEYLASLFPDSLIVKGFNVVSAMALQGPKDASRQVYICSNNIQARQVIE 180
 QY 181 LARQNFIPIDIGSLSSAREIENLPLRLFTLMRGPVVVAISLATFFFLYSFVRDVIHPYA 240
 DB 181 LARQNFIPIDIGSLSSAREIENLPLRLFTLMRGPVVVAISLATFFFLYSFVRDVIHPYA 240
 QY 241 RNQOSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTYRRFPFWLETWLQ 300
 DB 241 RNQOSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTYRRFPFWLETWLQ 300
 QY 301 CRKQGLLSFFFAWVAVASLCLPMRRSERYFLNMAVYQOVHANIENSWNEEVEWRIEMY 360
 DB 301 CRKQGLLSFFFAWVAVASLCLPMRRSERYFLNMAVYQOVHANIENSWNEEVEWRIEMY 360
 QY 361 ISFGIMSLGLSLAIVTSIPSVSNALNWRREFSIQSTLGYVALLISTFHVLIYGMKRAFE 420
 DB 361 ISFGIMSLGLSLAIVTSIPSVSNALNWRREFSIQSTLGYVALLISTFHVLIYGMKRAFE 420
 QY 421 BEYYRFTPPNPFVLAALVPSIVILDLQLCRYPD 454
 DB 421 BEYYRFTPPNPFVLAALVPSIVILDLQLCRYPD 454

RESULT 15
 US-10-455-822-121
 Sequence 121, Application US/10455822
 Publication No. US20040048798A1
 GENERAL INFORMATION:
 APPLICANT: Agensys, Inc.
 APPLICANT: Raitano, Arthur B.
 APPLICANT: Ge, Wangmao
 APPLICANT: Jakobovits, Aya
 APPLICANT: Chalilita-Eid, Pia M.
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer
 FILE REFERENCE: 51158-20016.24
 CURRENT APPLICATION NUMBER: US/10/455,822
 PRIOR FILING DATE: 2003-06-04
 PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
 PRIOR FILING DATE: 2002-12-20
 PRIOR APPLICATION NUMBER: US60/370,387
 PRIOR FILING DATE: 2002-04-05

```

; PRIOR APPLICATION NUMBER: US60/317, 840
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: US60/296, 656
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US60/091, 183
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US60/087, 520
; PRIOR FILING DATE: 1998-06-01
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 121
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-455-822-121

```

Query Match	100.0%;	Score 2351;	DB 12;	Length 454;
Best Local Similarity	100.0%;	Pred. No. 6.1e-222;		
Matches 454;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	MESISMGSPKSLSETCLPNGINGIKAKRVTVGIVIGSGDFAKSLTIRLIRCGYHVIGS	60
Db	1	MESISMGSPKSLSETCLPNGINGIKAKRVTVGIVIGSGDFAKSLTIRLIRCGYHVIGS	60
QY	61	RNPKEASEFFPHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVGKILIDVSNM	120
Db	61	RNPKEASEFFPHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVGKILIDVSNM	120
QY	121	RINQYESNAEYLAFLPDSLIVKGFNVASAMALQGPKDasRQVYICSNNIQARQVIE	180
Db	121	RINQYESNAEYLAFLPDSLIVKGFNVASAMALQGPKDasRQVYICSNNIQARQVIE	180
QY	181	LARQINFIPIDLGLSSAREIENLPLRLFTLWGPVVVAISLATEFFLYSFYEDVIHPYA	240
Db	181	LARQINFIPIDLGLSSAREIENLPLRLFTLWGPVVVAISLATEFFLYSFYEDVIHPYA	240
QY	241	RNQGSDEYKIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYYGTKYRRFPBWLETLQ	300
Db	241	RNQGSDEYKIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYYGTKYRRFPBWLETLQ	300
QY	301	CRKQLGLLSFFFAMVHVAYSLCLPMRRSERYLFELMAYQOVHANIENTSWNEEYVRIEMY	360
Db	301	CRKQLGLLSFFFAMVHVAYSLCLPMRRSERYLFELMAYQOVHANIENTSWNEEYVRIEMY	360
QY	361	ISFGIMSLGLISLLAVTSIPSVSNALNWRFEFSIOSTLGYVALLISTFHVLITYGKRAFE	420
Db	361	ISFGIMSLGLISLLAVTSIPSVSNALNWRFEFSIOSTLGYVALLISTFHVLITYGKRAFE	420
QY	421	EEYRFEYTPPNFVLALVLPISIVILDLQLCRYPD	454
Db	421	EEYRFEYTPPNFVLALVLPISIVILDLQLCRYPD	454

Search completed: April 16, 2004, 12:46:36
Job time : 49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 12:38:02 ; Search time 22 Seconds
(without alignments)
1065.372 Million cell updates/sec

Title: US-09-455-486-6
Perfect score: 2351
Sequence: 1 MESISMMSPKSLSETCLPN.....ALVLPSTIVLDLQLCRYPD 454

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	901	38.3	173	US-09-323-873A-8	Sequence 8, Appli
2	736	31.3	141	US-09-083-521-1	Sequence 1, Appli
3	717	30.5	339	US-09-323-873A-2	Sequence 2, Appli
4	717	30.5	339	US-09-685-166A-879	Sequence 879, App
5	144.5	6.1	227	US-09-655-270A-15	Sequence 15, Appl
6	144.5	6.1	227	US-09-955-597-17	Sequence 17, Appl
7	144.5	6.1	227	US-09-955-597-17	Sequence 17, Appl
8	107.5	4.6	695	US-08-487-886-2	Sequence 2, Appli
9	107.5	4.6	695	US-08-482-855-2	Sequence 2, Appli
10	107.5	4.6	695	US-08-474-986-2	GENERAL INFORMA
11	107	4.6	34	US-09-323-873A-20	Sequence 20, Appl
12	102.5	4.4	940	US-09-328-352-8165	Sequence 8165, Ap
13	101	4.3	476	US-09-316-083-3	Sequence 3, Appli
14	101	4.3	476	US-09-933-700-3	Sequence 3, Appli
15	101	4.3	724	US-09-252-991A-21494	Sequence 21494, A
16	99	4.2	365	US-09-170-496D-118	Sequence 118, App
17	99	4.2	365	US-09-170-496D-118	Sequence 226, App
18	99	4.2	365	US-09-364-425B-27	Sequence 2, Appli
19	99	4.2	531	US-08-724-974A-2	Sequence 2, Appli
20	97.5	4.1	390	US-08-460-576-2	Sequence 2, Appli
21	97	4.1	692	US-07-757-342D-6	Sequence 6, Appli
22	97	4.1	692	US-09-461-657B-6	Sequence 6, Appli
23	96	4.1	299	US-09-903-456-61	Sequence 61, Appl
24	96	4.1	1309	US-09-975-413A-10	Sequence 10, Appl
25	94.5	4.0	288	US-09-107-532A-6009	Sequence 6009, Ap
26	94.5	4.0	296	US-09-540-236-2922	Sequence 2922, Ap
27	94.5	4.0	345	US-09-489-039A-10740	Sequence 10740, A

28	94	4.0	284	4	US-09-903-456-83	Sequence 83, Appl
29	94	4.0	292	4	US-09-903-456-35	Sequence 35, Appl
30	94	4.0	293	4	US-09-903-456-45	Sequence 45, Appl
31	93	4.0	365	2	US-08-724-974A-3	Sequence 3, Appli
32	93	4.0	498	4	US-09-107-532A-7077	Sequence 7077, Ap
33	92.5	3.9	405	4	US-09-489-039A-9411	Sequence 9411, Ap
34	91.5	3.9	407	4	US-09-328-352-5605	Sequence 5605, Ap
35	91.5	3.9	495	4	US-08-985-343-2	Sequence 2, Appli
36	91	3.9	211	4	US-09-252-991A-23822	Sequence 23822, A
37	91	3.9	440	4	US-09-634-238-281	Sequence 281, App
38	91	3.9	550	4	US-09-543-681A-6026	Sequence 6026, Ap
39	91	3.9	808	4	US-09-134-001C-3105	Sequence 3105, Ap
40	90.5	3.8	340	4	US-09-543-681A-7503	Sequence 7503, Ap
41	90.5	3.8	388	4	US-09-107-532A-6381	Sequence 6381, Ap
42	90	3.8	500	4	US-09-178-093B-26	Sequence 26, Appli
43	90	3.8	970	4	US-09-795-927-7	Sequence 7, Appli
44	89.5	3.8	495	1	US-07-841-997A-2	Sequence 2, Appli
45	89.5	3.8	495	1	US-08-290-301-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-323-873A-8
; Sequence 8, Application US/09323873A
; Patent No. 6329503
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Kahan Leong
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Steve Chappell Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; FILE REFERENCE: 129.16USU2
; CURRENT APPLICATION NUMBER: US/09/323, 873A
; CURRENT FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087, 520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091, 183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-323-873A-8

Query Match 38.3%; Score 901; DB 4; Length 173;
Best Local Similarity 100.0%; Pred. No. 1.4e-86;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 246 DFKYPIPIVNTKLPVIAITLTLVLAAGLLAAAYQLYGTYRFRFPWLETWLCGRKQ 305
DB 1 DFKYPIPIVNTKLPVIAITLTLVLAAGLLAAAYQLYGTYRFRFPWLETWLCGRKQ 60
QY 306 GLLSFFPAMVHVAYSCLPMRSEERYLFLNMAYQOVHANIENSMNBEVWRIEMYSISFGI 365
DB 61 GLLSFFPAMVHVAYSCLPMRSEERYLFLNMAYQOVHANIENSMNBEVWRIEMYSISFGI 120
QY 366 MSLGLSLAVTSIPSVSNALNWRBSFIQSTLGYVALLISTFHYLIYGMKRA 418
DB 121 MSLGLSLAVTSIPSVSNALNWRBSFIQSTLGYVALLISTFHYLIYGMKRA 173

RESULT 2
US-09-083-521-1
; Sequence 1, Application US/09083521
; Patent No. 6048970
; GENERAL INFORMATION:

```

; APPLICANT: Lal, Preeti
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: PROSTATE GROWTH-ASSOCIATED MEMBRANE PROTEINS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/083,521
; FILING DATE: Herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0527 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 141 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSTUT10
; CLONE: 1691243
; US-09-083-521-1

Query Match      31.3%; Score 736; DB 3; Length 141;
Best Local Similarity 100.0%; Pred. No. 2.2e-69;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 314 MVHVAVSLCLPMRSEERYLFLNMAVOOVHANIEENSWNEEYRIEMYSISGINSGLSL 373
Db 1 MVHVAVSLCLPMRSEERYLFLNMAVOOVHANIEENSWNEEYRIEMYSISGINSGLSL 60

QY 374 LAVTSIPSVNALNWRSEFSFIQSTIGYVALLISTFHVLIYGMKRAPEEYRPTPNFV 433
Db 61 LAVTSIPSVNALNWRSEFSFIQSTIGYVALLISTFHVLIYGMKRAPEEYRPTPNFV 120

QY 434 LALVLPSTIVLDLQLCRYPD 454
Db 121 LALVLPSTIVLDLQLCRYPD 141

RESULT 3
US-09-323-873A-2
; Sequence 2, Application US/09323873A
; Patent No. 6329503
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Kahan Leong
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Steve Chappell Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
; FILE REFERENCE: 129.16US02
; CURRENT APPLICATION NUMBER: US/09/323,873A
; CURRENT FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520

```

```

; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-09-323-873A-2

Query Match      30.5%; Score 717; DB 4; Length 339;
Best Local Similarity 54.9%; Pred. No. 8.5e-67;
Matches 130; Conservative 48; Mismatches 59; Indels 0; Gaps 0;

QY 208 LFTLMRGPVVVAISLATFFFLYSFVDVHHPYARNOQSDPYKIPPIEIVNKTLPVATLL 267
Db 67 LFPQWHLPIKIAIIASLTFLLYTLREVIHPLATSHQQPYKIPPIVINKVLPVVSITLL 126

QY 268 SLVYLAGLLAAAYQLYGTGTRRRPPLWLTWQCRKQGLISFFPAMHVAVSLCLPMR 327
Db 127 ALVYLPGVIAIVQLHNGTKYKKPFPHLDKMLTRKQGLISFFPVLHVAISLSYPMR 186

QY 328 SERVFLNMAVOOVHANIEENSWNEEYRIEMYSISGINSGLSLAVTSIPSVNALN 387
Db 187 SYRYKLINMAVOOVQONKEDAWIEHDVWRMEIYVSLGIVGLAIALAVTSIPSVSDSLT 246

QY 388 WREFSFIQSTIGYVALLISTFHVLIYGMKRAPEEYRPTPNFVLALVLPSTIVL 444
Db 247 WREFHYIQSKLGIVSLGLTGHALITPAWKNWIDIKQFVWYTPPTFMIAVFLPIVLI 303

RESULT 4
US-09-685-166A-879
; Sequence 879, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, David C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Ronger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedrick, Thomas S.
; APPLICANT: Carter, Darrick L.
; APPLICANT: Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-685-166A-879

Query Match      30.5%; Score 717; DB 4; Length 339;
Best Local Similarity 54.9%; Pred. No. 8.5e-67;
Matches 130; Conservative 48; Mismatches 59; Indels 0; Gaps 0;

QY 208 LFTLMRGPVVVAISLATFFFLYSFVDVHHPYARNOQSDPYKIPPIEIVNKTLPVATLL 267

```



```
Db      67 LFPQWHLPIKIAAIIASLTFLYTLTREVIHPLATSHQQYFYKIPILVINKVLPMWSITLL 120
QY      268 SLVYLGLLAALAAAYQLYYGTYKRFRFPMLETWLOCRKQLGLSFFFAWVAAYSCLPMRR 327
        :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Dd      127 ALVYLPGVIALAVQLHNSTKYKKFPHMLDKMULTRKQFGLSFFFAVLHAISLSYPERR 186
QY      328 SERYLFLNMAVOOVHANIEINSWNBEZYRIEMYISFGIMSLGLLSLLAVTSIPSVSNALN 387
        |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Dd      187 SYRYKLINMAVOOVQNKEADAMIEHDVMREIIVYSLGIVGAILALLAVTSIPSVSDSLT 246
QY      388 WREFSFIOSTLGYVALLIISTFHVLITYGWKRAFEEBEYRFTYPPNFVLALVLPISVIL 444
        |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Dd      247 WRPFHYIOSKLGIVSLIGTIHALIFANMKMIDIKQFVWYTPTPRTMIAVFLPIVAVLI 303
```

RESULT 5

```

US-09-655-270A 15
; Sequence 15, Application US/09655270A
; Patent No. 6329151
;
; GENERAL INFORMATION:
; APPLICANT: Rouviere, Pierre B.
; TITLE OF INVENTION: High Density Sampling of Differentially Expressed Prokaryotic mRNAs
; FILE REFERENCE: BC1011 US NA
; CURRENT APPLICATION NUMBER: US/09/655,270A
; CURRENT FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/120,702
; PRIOR FILING DATE: 1999-February-19
; PRIOR APPLICATION NUMBER: 60/152,542
; PRIOR FILING DATE: 1999-September-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
;
; LENGTH: 227
;
; TYPE: PRT
;
; ORGANISM: Rhodococcus erythropolis HL PM-1
US-09-655-270A-15

```

Query Match	6.18;	Score 144.5;	DB 4;	Length 227;
Best Local Similarity	28.0%;	Pred. No. 6.4e-07;		
Matches 60;	Conservative 33;	Mismatches 88;	Indels 33;	Gaps 9;

```

QY      25 IKDARKVTGVTGSGDFAKSLTIRLIRCGYHVVIGSRNPKFAGEEPH-----VVD 75
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1 MKSSKIAVVG--GTGPGQKGLAYRFAAGWVVIIGSRAERAEEALAEVRRAGDGA VVS 58
QY      76 VTHHEDALTKNTNIIFVAI----HREHYTSLMDLRHLVKGILL-----DVSNNMRIN 123
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      59 AADNASAAADCPRIILLVVPYDGHRELVS---ELAFIPAGKLVVSCVNPGLGFDKSGAYGID 115
QY      124 QYPESNAEYVLASLPEDSLIVKGFNVVSA---WALQLGPKDASROYICSNNIQARQVIE 180
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      116 VEEGSAAEQLRDLVPGATVVAAFHHLTSAVNIWEHE-GP--LPEDVLVCGDDRSKADDEVAR 172
QY      181 LARQLNFIIP-IDIGSLSSAREIENLPRLIFTLMR 213
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      173 LAVAITGRPGIDGALRVARQLEPPLTAVLINVR 206

```

RESULT 6

```

US-09-651,941-17
; Sequence 17, Application US/09651941
; Patent No. 6355470
; GENERAL INFORMATION:
; APPLICANT: ROUVIER, PIERRE E
; APPLICANT: WALTERS, DANA M
; APPLICANT: RAINER, RUSS
; TITLE OF INVENTION: Genes Encoding Picric Acid Degradation
; FILE REFERENCE: BC1022 US NA
; CURRENT APPLICATION NUMBER: US/09/651, 941
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/152,545
; PRIOR FILING DATE: 1999-10-03
; NUMBER OF SEQ ID NOS: 28

```

```

; SOFTWARE: Microsoft Office 97
; SEQ ID NO: 17
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Rhodococcus erythropolis HL PM-1
US-03-651-941-17

```

Query Match

Query Match	6.14;	Score 144.5;	DB 4;	Length 227;
Best Local Similarity	28.0%;	Pred. No. 6.4e-07;		
Matches 60;	Conservative 33;	Mismatches 88;	Indels 33;	Gaps 9;

OY 25 IKDARKVTGVIGSGDFAKSLTRLRLRCGHVIVIGSRNPKFASEFFPH-----VD 75
 | : | : | : ||| : :
DB 1 MMSKIAVVG--GTGPQGKLAYRFAAGMPVTVGSRSAERAEAALEVRRRAGDAVS 58

QY 76 VTHEDALTKNIIVAI---HREHYTSLWDLRHLVGKILI-----DVSNNMRIN 123

Db 59 AADNAAADCPILLVVPYDGHRELV--ELAPIFAGKLVSQVNPGLGFDKSGAYGLD 115

124 QYPESNAEYLASLFPDSLIVKGFNVSA--WALQLGPKDASRQVYICSNNIQARQVIE 180

Db 116 VEEGSAEQLRDLVPGATVVAFFHLSAVNLWEHE-GP--LPEDVLVCGDDRSKDEVAR 172

QY 181 LARQNFIP-IDGSSAREIENLPRLFTLWR 213

Db 173 LAVAITGRPGIDGALRVARQLEPLTAVLINVR 206

RESULT 7

```

US-09-955-597-17
? Sequence 17, Application US/09955597
? Patent No. 6461856
? GENERAL INFORMATION:
? APPLICANT: ROUVIER, PIERRE F
? APPLICANT: WALTERS, DANA M
? APPLICANT: RAINER, RUSS
? TITLE OF INVENTION: Genes Encoding Picric Acid Degradation
? FILE REFERENCE: BC1022 US NA
? CURRENT APPLICATION NUMBER: US/09/955,597
? PRIOR FILING DATE: 2001-09-17
? PRIOR APPLICATION NUMBER: 60/152,545
? NUMBER OF SEQ ID NOS: 28
? SOFTWARE: Microsoft Office 97
? SEQ ID NO 17
? LENGTH: 227
? TYPE: PRT
? ORGANISM: Rhodococcus erythropolis HL PM-1
US-09-955-597-17

```

Query Match	6.18;	Score 244.5;	DB 4;	Length 227;
Best Local Similarity	28.0%;	Pred. No. 6.4e-07;		
Matches 60;	Conservative 33;	Mismatches 88;	Indels 33;	Gaps 9;

QY 25 IKDARKVTGVIIGSGDPAKSLTIRLLRCGYHVIIGSRNPKFASEFFPH-----VVD 75

Db 1 MKSKIAVVG--GTGPQKGGLAYRFAAGWPVVIIGSRSAERAEAEALEVRRRAGDGAVVS 58

76 VTHHEDALTKINIIFVAI---HREHYTSLWDLRHLVGKILI-----DVSNNMRIN 123

Db 59 AADNASAAADCPILLVVPYDGHRELVS--ELAPIFAGKLWVSCVNP LGFDKSGAYGLD 115

124 QYPESNAEYIASLFPDSLIVKGFNVSA--WALQLGPKDASRQVICSNNIQARQVIE 180

Db 116 VEEGSAEQLRDLVPGATVVAAPHLSAVNLWEHE-GR--LPEDVLVCGDDRSAXDEVAR 172

QY 181 LARQNFIP-IDLGLSSAREIENPLRLFTLWR 213

Db 173 LAVAITGRPGIDGGALRVARQLEPTAVLINNR 206

9
8
7
6
5
4
3
2
1
E
C

US-08-48

US-08-487-886-2

Sequence 2, Application US/08487886
Patent No. 5744448
GENERAL INFORMATION:
APPLICANT: Kelton, Christie Ann
APPLICANT: Schweickhardt, Rene Lynn
APPLICANT: Cheng, Shirley Vui Yen
APPLICANT: Nugent, No. 5744448een Patrice
TITLE OF INVENTION: Human Follicle Stimulating
TITLE OF INVENTION: Hormone Receptor
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Stephan P. Williams,
ADDRESSEE: Ares-Serono, Inc.
STREET: Exchange Place, 37th floor
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
COMPUTER: IBM PS/2, model 55 SX
OPERATING SYSTEM: MS-DOS version 4.0
SOFTWARE: VAX/VMS Massail via Kermit to IBM MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,886
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/670,085
FILING DATE: 15-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Stephan P.
REGISTRATION NUMBER: 28546
REFERENCE/DOCKET NUMBER: US/252
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 723-1300
TELEFAX: (617) 723-8923
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 695
TYPE: Amino acid
TOPOLOGY: Linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: signal sequence
LOCATION: -17 to -1
IDENTIFICATION METHOD: hydrophobic
FEATURE:
NAME/KEY: putative amino-terminal extracellular domain
LOCATION: 1 to 349
IDENTIFICATION METHOD: similarity with other
IDENTIFICATION METHOD: dimeric glycoprotein receptor extracellular
IDENTIFICATION METHOD: domains, hydrophilic
FEATURE:
NAME/KEY: transmembrane domain
LOCATION: 350 to 613
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane domains
FEATURE:
NAME/KEY: putative transmembrane region I
LOCATION: 350 to 370
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region II
LOCATION: 382 to 404
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region III
LOCATION: 427 to 448

IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region IV
LOCATION: 469 to 491
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region V
LOCATION: 512 to 533
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region VI
LOCATION: 557 to 580
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region VII
LOCATION: 592 to 613
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative carboxy-terminal intracellular
NAME/KEY: domain
LOCATION: 614 to 678
US-08-487-886-2
Query Match 4.6%; Score 107.5; DB 1; Length 695;
Best Local Similarity 17.0%; Pred. No. 0.028;
Matches 84; Conservative 77; Mismatches 148; Indels 185; Gaps 17;
QY 79 HEDALTKTNIIIFVAIRHREHYTSLMDRLHVG-----KILIDVSNMR 121
DB 98 HEIRIEKANNL-LYINBEAFQNLPLQYLLISNTGKHLPDVHKHSLQKVLDDIQNVN 156
QY 122 INQYBESNAEYLASLTFPSLIVKGFNVSAWALQGPKASROYVICSNNIOARQVTEL 181
DB 157 IH-----TIERSNFVGLSFESVILML-----NKGIOEIHNCA----- 189
QY 182 ARQNLPIPLDGLSLSAREIENLPLRLFTLRGPVVAISLAPFFELYSFVRDVHPYAR 241
DB 190 ---FNGTQDELNLSDNNLLELPNDVFGASGPVILDISRTIHSLSYGLNLKCLR 246
QY 242 NQGSDFYKPIPIEIVNKL-PIVAITLTLSTV-----PWLSTWLQCRQL 271
DB 247 RSTYNLKLIP-----TLEKLVALMEASLTPSHCCAFAWNRROISELHPICNKSILQE 300
QY 272 -----LAGLAAAYQLYYGKRRF-----PWLSTWLQCRQL 305
DB 301 VDYNTQTRGQRSSLAEDNESSYSRGFDWYTEFDYDLNCEVVVDYTCSPKDAFNPCEIDIM 360
QY 306 G-----LISFFFAWVHA-----YSLCLPMRRSERLYFLMAYQQVHANEN 347
DB 361 GYNILRLVLIWIFISILATIGNIIVLVITTSQYKLTVP-----RLMGNLAIFAD----- 408
QY 348 SWNEBEVWRIMYISFGINSGLSLAVTSIPSYSNALNRE-----PSFIQSTL 398
DB 409 -----LCIGYVLLIASVDIHTKSQYHNYAIDWQTAGCDAAAGFTVPASEL 455
QY 399 G---VYALLISTFH-----VLITGKRAPEEYRYRFTPPNFV-LA 435
DB 456 SVYTLTALTLEWHITTHAMQDDCKQVLRHAASVWVGWIFAFRAALFPPIFGISSYMKYS 515
QY 436 LVLPISVILDLQL 449
DB 516 ICLPMDIDPLSQL 529

RESULT 9
US-08-482-855-2
Sequence 2, Application US/08482855
Patent No. 6121016
GENERAL INFORMATION:
APPLICANT: Kelton, Christie Ann
APPLICANT: Schweickhardt, Rene Lynn
APPLICANT: Cheng, Shirley Vui Yen
APPLICANT: Nugent, No. 6121016een Patrice
TITLE OF INVENTION: Human Follicle Stimulating
TITLE OF INVENTION: Hormone Receptor
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Stephan P. Williams,
ADDRESSEE: Ares-Serono, Inc.
STREET: Exchange Place, 37th floor
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
COMPUTER: IBM PS/2, model 55 SX
OPERATING SYSTEM: MS-DOS version 4.0
SOFTWARE: VAX/VMS Mas811 via Kermit to IBM MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,855
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/670,085
FILING DATE: 15-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Stephan P.
REGISTRATION NUMBER: 28546
REFERENCE/DOCKET NUMBER: US/252
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 723-1300
TELEFAX: (617) 723-8923
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 695
TYPE: Amino acid
TOPOLOGY: Linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: signal sequence
LOCATION: -17 to -1
IDENTIFICATION METHOD: hydrophobic
FEATURE:
NAME/KEY: putative amino-terminal extracellular domain
LOCATION: 1 to 349
IDENTIFICATION METHOD: similarity with other
IDENTIFICATION METHOD: dimeric glycoprotein receptor extracellular
IDENTIFICATION METHOD: domains, hydrophilic
FEATURE:
NAME/KEY: transmembrane domain
LOCATION: 350 to 613
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane domains
FEATURE:
NAME/KEY: putative transmembrane region I
LOCATION: 350 to 370
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region II
LOCATION: 382 to 404
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length

FEATURE:
NAME/KEY: putative transmembrane region III
LOCATION: 427 to 448
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region IV
LOCATION: 469 to 491
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region V
LOCATION: 512 to 533
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region VI
LOCATION: 557 to 580
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region VII
LOCATION: 592 to 613
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative carboxy-terminal intracellular
LOCATION: 614 to 678
US-08-482-855-2

Query Match 4.6%; Score 107.5; DB 3; Length 695;
Best Local Similarity 17.0%; Pred. No. 0.028;
Matches 84; Conservative 77; Mismatches 148; Indels 185; Gaps 17;
QY 79 HEDALTKNTIIFVAIHREHYTSLMDLHLVVG-----KILIDVSNMR 121
DB 98 HEIRIEKANNL-LYINPEAFQNLLENLYLLISNTGKHLPDVHKHSLQKVLDDIQDNIN 156
QY 122 INQYPSNAEYLASLPDLSLYKGNVVSAMALQGPXDASROVYICSNNIQARQVIEL 181
DB 157 IH-----TIERNSFVGLSFESVITLW-----NKGIGQEIHNCA----- 189
QY 182 ARQNFIPIDGLSSAREIENPLRLFTLMRGPVVVAISLATFFLYSFVRDYIHPYAR 241
DB 190 ---FNGTQDLDELNTSDNNLELNDVDFHAGSPVILDISRTRHSLPSYGLNKLRA 246
QY 242 NQOSDFYKIPLEIVNKTLP-PIVATLLSLVY-----PWLFTWLCGRKQL 305
DB 247 RSTYNLKKLP-----TLEKLVLMASLTYPHSCAFANRRQISELHPICNSILRQE 300
QY 272 -----LAGLLAAAYQLYGTGKRRF-----PWLFTWLCGRKQL 305
DB 301 VDMYQTRGROSSLAEDNESSYSGFDMTYTEFDYDLCNEVVDTCSFKPDAPNCEIDIM 360
QY 306 G-----LTSFFFAWVHA-----YSICLPWRRSERYLFLNMAQQVHANIEN 347
DB 361 GYNIRVLWIFISILAITGNIIVILLTSQYKLTVP-----RFLMCNLAFAD----- 408
QY 348 SMNEBEVVRIMYISFGIMSLGLSLAVTSIPSVSNALNRE-----PSFIQSTL 398
DB 409 -----LCIGIYLLIASVDIHTKSQYHNVAIDWQTGACDAGGFTVFASEL 455
QY 399 G---YVALLISGFH-----VLIYGNKRAFEERYRFTYPPNFV-LA 435
DB 456 SVYTLTATLTERWHTITHAMQLDKVQLRHAASVVMGMWIFAFAALFPIFGISSYMKVS 515
QY 436 LVLPSIVILDLIQL 449

Db 516 ICLPMDIDSPLSQL 529

RESULT 10

US-08-474-986-2

GENERAL INFORMATION:

APPLICANT: Kelton, Christie Ann

Schweickhardt, Rene Lynn

Cheng, Shirley Yui Yen

Nugent, No. 637271leen Patrice

TITLE OF INVENTION: Human Follicle Stimulating Hormone Receptor

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Stephan P. Williams,

STREET: Exchange Place, 37th floor

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density

COMPUTER: IBM PS/2, model 55 SX

OPERATING SYSTEM: MS-DOS version 4.0

SOFTWARE: VAX/VMS Mass11 via Kermit to IBM MS-DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/474, 986

FILING DATE: 07-Jun-1995

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/670, 085

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Williams, Stephan P.

REGISTRATION NUMBER: 28546

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 723-1300

TELEFAX: (617) 723-8923

LOCATION: 614 to 678

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-08-474-986-2

Query Match 4.6%; Score 107.5; DB 4; Length 695;

Best Local Similarity 17.0%; Pred. No. 0.028;

Matches 84; Conservative 77; Mismatches 148; Indels 185; Gaps 17;

Db 79 HEDALTKNITFVAIHREHYTSLWDLHLLVG-----KILIDVSNMR 121

98 HEIRIEKANL-LYINPEAFQNLPLNQLLISNTGKILPDVHKIHSLOKVLDDIQDNIN 156

QY 122 INQPEESNAEYLASLFPDSLIVGFNVSAWALQLGPKASQVYICSNINQARQVIEL 181

157 IH-----TIERNSFVGLSFESVILWL-----NKNGIQEIHNC-----189

QY 182 ARQLNFIPIDGLSSAREIENPLRFTLWRGPVVAISLATFPFLYSFVRDVIHPYAR 241

190 ---FNGTQDELNLSNNELELPNDVHGASGPVILDIRTRHSLPSYGLNKKLRA 246

QY 242 NQSDFYKPIETVNTL-PIVATILSLVY-----271

247 RSTYNLKKLP-----TLEKVALMEASLTPSHCCAFANWRQISELHPICNKSILRQE 300

QY 272 -----LAGLLAAAYQLYGTXYRRF-----PPWLETWLOCRKQL 305

Db 301 VDYMTQTRGRRSLAEDNESSYSRGFDWYTEFDYDLCEVVDVTCSPKPAFNCEDIM 360

QY 306 G-----LISFFAMVAVY-----YSLCLPMRSEERYLFLNMAYQOVHANIN 347

361 GYNILRVLIWIFISILATIGNIIVLVILTTSQYKLTVP-----RFLMCNLAFAF-----408

QY 348 SWNEEVRWRIEMYSFGIMSLGLSLAVTSIPSVSNALNMR-----FSFIQSTL 398

Db 409 -----LCIGIYLLIASVDIHTKSQYHNVAIDWQTGAGCDAAEFFTVASEL 455

QY 399 G---YVALLISTPH-----VLIYGMKRAFEERYRYTPPNFV-LA 435

Db 456 SVYTLTATIERWHITTHAMQLDCKVQLRHAASVWVWGMIFAFPAALFPITGISSYMKVS 515

QY 436 LVLPSTVILDLQL 449

Db 516 ICLPMDIDSPLSQL 529

RESULT 11

US-09-323-873A-20

Sequence 20, Application US/09323873A

Patent No. 6329503

GENERAL INFORMATION:

APPLICANT: Daniel E. Afar

APPLICANT: Rene S. Hubert

APPLICANT: Kahan Leong

APPLICANT: Arthur B. Raitano

APPLICANT: Douglas C. Saffran

APPLICANT: Steve Chappell Mitchell

TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS

FILE REFERENCE: 129.16USU2

CURRENT APPLICATION NUMBER: US/09/323, 873A

CURRENT FILING DATE: 1999-06-01

PRIOR APPLICATION NUMBER: 60/087,520

PRIOR FILING DATE: 1998-06-01

PRIOR APPLICATION NUMBER: 60/091,183

PRIOR FILING DATE: 1998-06-30

NUMBER OF SEQ ID NOS: 32

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 20

LENGTH: 34

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: STEAP-1 PEPTIDE

US-09-323-873A-20

Query Match 4.6%; Score 107; DB 4; Length 34;

Best Local Similarity 58.8%; Pred. No. 0.0003;

Matches 20; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Db 326 RRSERYFLNVAAYQOVHANINENSWNEEVRWRIEM 359

1 RRSRYKLLNVAAYQOVQNKEDAVIEHDVWRMEI 34

RESULT 12

US-09-328-352-8165

Sequence 8165, Application US/09328352

Patent No. 6562958

GENERAL INFORMATION:

APPLICANT: Gary L. Breton et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

FILE REFERENCE: GTC99-03PA

CURRENT APPLICATION NUMBER: US/09/328, 352

CURRENT FILING DATE: 1999-06-04

NUMBER OF SEQ ID NOS: 8252

SEQ ID NO 8165

LENGTH: 940

TYPE: PRT

ORGANISM: Acinetobacter baumannii

US-09-328-352-8165

Query Match 4.4%; Score 102.5; DB 4; Length 940;

Best Local Similarity 19.5%; Pred. No. 0.15;

Matches 71; Conservative 74; Mismatches 146; Indels 73; Gaps 17;


```
QY      88 IIFVAIHREHYTSLMDLRHLVGLIDVSNMNRINQYPESNAEYIASL-----FPDSLI 142
      262 VVFLTLVESHFGV--ITSLTTASILEFTILLAITVFLSLKQQAITYIALALGMAYAAPLV 319
Db      143 VKGF--NVVSAMALQLGPKDASRQVYICSNNIQARQOVIELARQLN----FIPIDLSLS 196
      320 IPQYRPDVVFLFS-----YYLVINLAVAVNFIQPKWILNQIAFFATMFIGGSA 368
QY      197 SAREIENLPLRLFTL-WRGPVVVAISLATFFFLYSFVRDVIHPYARNQSDFYKIPIEIV 255
      369 IAFYAE--PAKFDTLDM---ILMLHIALFIMLSVRYSQNTISRVSHEKQEGIRLP--- 418
QY      256 NKTLPVAVITLLSLVYLAGLLAAAYQLYGTGK-----YRRPPWLETWLQCRK 303
      419 ---PLLDVGLIFNVFVLGFTLHAYLVHESQTALTIGAAVLAGTYAVLTFWIK--KTHP 471
Db      304 QLGLSFFPAMVHA-YSLCLPMRSERYLFLNMAYQQVHANIEGNEEVEWRIEMYIS 362
      472 QLSVLAKSFFILAVAFALIFPLAKGAHWTAIGWVAQGTALIV--WGYTERYRLSRYI- 527
QY      363 FGIMSLGLSLAAVTSIPSVSNALNW--REFSFIQSTLGYVALLISTFHYLIYWKRAFE 420
      528 -GVILVLLSSIALFYQV-----WANEFPILTSTSIYIAQFISAFYLLQYNSK---E 575
Db      421 EEEY 424
QY      576 QRYF 579

RESULT 13
US-09-316-083-3
; Sequence 3, Application US/09316083A
; Patent No. 6280942
; GENERAL INFORMATION:
; APPLICANT: The Institute of Physical and Chemical Research
; TITLE OF INVENTION: Endonuclease
; FILE REFERENCE: PH-651
; CURRENT APPLICATION NUMBER: US/09/316, 083A
; EARLIER FILING DATE: 1999-05-20
; EARLIER APPLICATION NUMBER: JP98/141861
; EARLIER FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; US-09-316-083-3

Query Match      4.3%; Score 101; DB 3; Length 476;
Best Local Similarity 20.7%; Pred. No. 0.075;
Matches 52; Conservative 48; Mismatches 85; Indels 66; Gaps 12;
```

```
QY      62 NPKPASEFPVHVDTVTHEDALTKTNI---IFVAIHREHYT-----SLMD 103
      254 NPYFVNAFSINI-----KTNLAKEKIFNTIYNKLYSDYKINQINNHI PYNNYLK 302
Db      104 LRHLVGLKILIDVSNMNRINQYPESNAEYIASLF--PDSLIVKGFNVVSAMALQLGPKDA 161
      303 INNKLPIKINMIDIKNNYWLAGFTAADGSLSSMYPNPKDTLLFKNM----- 347
QY      162 SRQYICSNNIQARQOVIELARQLNFIPIDLSLSAREIENLPLRLFTLWRGPVVVAIS 221
      348 -RPSYVIS-QVETRKELIYLIG-----SFDL-SISNVKVGNRKLDKFKLFRTTDELMMK 400
Db      222 LATFFFLYSFVRDVIHPYARNQSDFYKIP---IEIVNKTLPVAVITLLSLVYLAGLLA 277
      401 -----FIYF--DKFLPLHDNKQFNYIKFRNTFIKSYNMNRRVFGVLVSE--YINNIKI 451
QY      278 AAYQLYGTGY 288
Db      452 DNYDYYYNKKY 462
```

```
RESULT 14
US-09-933-700-3
; Sequence 3, Application US/09933700
; Patent No. 6528296
; GENERAL INFORMATION:
; APPLICANT: The Institute of Physical and Chemical Research
; TITLE OF INVENTION: Endonuclease
; FILE REFERENCE: PH-651
; CURRENT APPLICATION NUMBER: US/09/933, 700
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 09/316, 083
; PRIOR FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; US-09-933-700-3

Query Match      4.3%; Score 101; DB 4; Length 476;
Best Local Similarity 20.7%; Pred. No. 0.075;
Matches 52; Conservative 48; Mismatches 85; Indels 66; Gaps 12;
```

```
QY      62 NPKPASEFPVHVDTVTHEDALTKTNI---IFVAIHREHYT-----SLMD 103
      254 NPYFVNAFSINI-----KTNLAKEKIFNTIYNKLYSDYKINQINNHI PYNNYLK 302
Db      104 LRHLVGLKILIDVSNMNRINQYPESNAEYIASLF--PDSLIVKGFNVVSAMALQLGPKDA 161
      303 INNKLPIKINMIDIKNNYWLAGFTAADGSLSSMYPNPKDTLLFKNM----- 347
QY      162 SRQYICSNNIQARQOVIELARQLNFIPIDLSLSAREIENLPLRLFTLWRGPVVVAIS 221
      348 -RPSYVIS-QVETRKELIYLIG-----SFDL-SISNVKVGNRKLDKFKLFRTTDELMMK 400
Db      222 LATFFFLYSFVRDVIHPYARNQSDFYKIP---IEIVNKTLPVAVITLLSLVYLAGLLA 277
      401 -----FIYF--DKFLPLHDNKQFNYIKFRNTFIKSYNMNRRVFGVLVSE--YINNIKI 451
QY      278 AAYQLYGTGY 288
Db      452 DNYDYYYNKKY 462
```

```
RESULT 15
US-09-252-991A-21494
; Sequence 21494, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252, 991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074, 788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094, 190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21494
; LENGTH: 724
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-21494

Query Match      4.3%; Score 101; DB 4; Length 724;
Best Local Similarity 20.5%; Pred. No. 0.14;
Matches 72; Conservative 57; Mismatches 105; Indels 118; Gaps 16;
```

QY 132 YLASFPSLIVKGFNVSAWALQIGPKDASRQVYICSNINQA-----ROQVIELARQIN 186
Db 298 YLSVVLPSDL-SKTLAMWMAVVLVCGTLESA---LCVLSLSLGGPHRQALDILRQA 352
QY 187 FIPIDL-GSLSSAREIENLPLRLFTLMRGPVVAISLA-----TFFFLYSFVRDVTHP 238
Db 353 FRPLWLIGSLALGEVAHDPRLIAGLGEHISICLSTLANASALFTALFVMRFRPIAH- 411
QY 239 YARNQC-----SDFYKIPi-----EIVNK 257
Db 412 LIRNQLERRLKRSLSHDLVQLVGSIMFVPLVLVGISLFATFVSAGDSSALRRALVCA 471
QY 258 TLPIVAITLLSLVYLAGLLAAAYQLYGTXYRRFPFWLETWLQCRKQLGLLSFFAMVHV 317
Db 472 VLAVVAMTVIGLI-----RRSSRVGAGPRRSAPYIEQ-----LQSGYTLIHI 515
QY 318 AYSLCLPMRRSERYLFLNMAAYQOVHANIENTSMNEEVMRIEM-YISFGIM----- 366
Db 516 FVVL-----FFIEVALRWGMSLIR-YABEGEQISMKVVSFGTLLVAVLWIL 564
QY 367 -----SLGL-----LSLLAVTSIPSVSNALNWRPFSFIQSTLGYVALLIS 406
Db 565 TDTAIOHSLGLGKSRPNTRALTMLPLIRNVL-----FATIAVIALIVA 608

Search completed: April 16, 2004, 12:41:47
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 12:32:36 ; Search time 60 Seconds
(without alignments)
2137.943 Million cell updates/sec

Title: US-09-455-486-6
Perfect score: 2351
Sequence: 1 MESISMGSPKSLSETCLN.....ALVLPISIVLDLQLCRYPD 454

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1:	Geneseqp1980s:*
2:	Geneseqp1990s:*
3:	Geneseqp2000s:*
4:	Geneseqp2001s:*
5:	Geneseqp2002s:*
6:	Geneseqp2003as:*
7:	Geneseqp2003bs:*
8:	Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2351	100.0	454	4	AAE02781	Aae02781 Human six
2	2351	100.0	454	4	AAU10188	Aau10188 Human ORF
3	2351	100.0	454	5	ABG61933	Abg61933 Prostate
4	2351	100.0	454	5	AAU76538	Aau76538 Tumour-as
5	2351	100.0	454	5	AAU80190	Aau80190 Human PUM
6	2351	100.0	454	7	ABU08893	Abu08893 Tumour-as
7	2348	99.9	454	7	ADB65001	Adb65001 Human pro
8	2294	97.6	490	5	AAE28951	Aae28951 Human STE
9	2294	97.6	490	7	ABU63312	Abu63312 Human tra
10	2290	97.4	490	4	AAU10187	Aau10187 Human six
11	2036	86.6	419	4	AAU10189	Aau10189 Human ORF
12	1972.5	83.9	576	4	ABG12306	Abg12306 Novel hum
13	1972.5	83.9	1273	4	ABG00113	Abg00113 Novel hum
14	1738	73.9	450	4	AAE02841	Aae02841 Human STE
15	1364.5	58.0	1082	5	ABP62883	Abp62883 Human pol
16	1316	56.0	488	4	AAB49483	Aab49483 Rat p-HYD
17	1316	56.0	526	5	ABB83365	Abb83365 Murine tu
18	1272	54.1	488	5	ABB85775	Abb85775 Human dru
19	1272	54.1	488	5	ABB83366	Abb83366 Human tum
20	1269	54.0	488	4	AAB93224	Aab93224 Human pro
21	1269	54.0	488	4	AAU04564	Aau04564 Human G-P
22	1269	54.0	488	4	AAU10220	Aau10220 Human six
23	1269	54.0	488	6	ABU60864	Abu60864 Human G P
24	1245.5	53.0	487	4	AAB49481	Aab49481 Human P-H
25	1088	46.3	459	4	AAB74715	Aab74715 Human mem

26	1087	46.2	459	4	AAE02782	Aae02782 Human six
27	1087	46.2	459	5	ABP64820	Abp64820 Human pro
28	1085	46.2	459	4	AAU10190	Aau10190 Human six
29	1082	46.0	458	4	AAE02636	Aae02636 Human STE
30	1074.5	45.7	456	4	AAB49482	Aab49482 Human STR
31	901	38.3	173	3	AAV58195	Aay58195 Human pro
32	844	35.9	179	4	AAU18006	Aam18006 Peptide #
33	844	35.9	179	4	ABB37041	Abb37041 Peptide #
34	844	35.9	179	4	AAU30517	Aam30517 Peptide #
35	844	35.9	179	4	ABB31810	Abb31810 Peptide #
36	844	35.9	179	4	ABB22356	Abb22356 Protein #
37	844	35.9	179	4	AAU70183	Aam70183 Human bon
38	844	35.9	179	4	AAU57768	Aam57768 Human bra
39	844	35.9	179	4	ABG51883	Abg51883 Human liv
40	844	35.9	179	4	AAU05646	Aam05646 Peptide #
41	844	35.9	179	5	ABG39817	Abg39817 Human pep
42	844	35.9	237	4	AAU04565	Aau04565 Human G-P
43	768	32.7	237	6	ABU60865	Abu60865 Human G P
44	736	31.3	141	3	AAV52589	Aay52589 Human pro
45	717	30.5	267	6	ABU60886	Abu60886 Human G P

ALIGNMENTS

RESULT 1			
AAE02781	ID	AAE02781	standard; protein; 454 AA.
XX	XX	AAE02781;	
AC	XX	06-AUG-2001	(first entry)
DT	XX	Human six transmembrane epithelial antigen of prostate (STEAP)-2 protein	
XX	XX	Human, cytostatic; antiproliferative; vaccine; gene therapy;	
KW	KW	six transmembrane epithelial antigen of the prostate-2; STEAP-2;	
KW	KW	chromosome 7q21; cancer; prostate; colon; bladder; lung; ovarian;	
KW	KW	pancreatic.	
XX	XX	Homo sapiens.	
OS	XX	Location/Qualifiers	
FH	FT	100. .108	/label= HLA-A2_binding_peptide #5
FT	FT	153. .165	/label= Immunogenic_peptide #1
FT	FT	207. .228	/label= Transmembrane_domain #1
FT	FT	227. .235	/label= HLA-A2_binding_peptide #1
FT	FT	255. .277	/label= Transmembrane_domain #2
FT	FT	304. .325	/label= Transmembrane_domain #3
FT	FT	306. .314	/label= HLA-A2_binding_peptide #4
FT	FT	307. .315	/label= HLA-A2_binding_peptide #3
FT	FT	345. .358	/label= Immunogenic_peptide #2
FT	FT	359. .381	/label= Transmembrane_domain #4
FT	FT	393. .415	/label= Transmembrane_domain #5
FT	FT	402. .410	/label= HLA-A2_binding_peptide #2
FT	FT	428. .450	/label= Transmembrane_domain #6
XX	XX	WO200140276-A2.	
PN	XX	07-JUN-2001.	
XX	XX		
PD	PD		

XX 06-DEC-2000; 2000WO-US033040.
PF
XX 06-DEC-1999; 99US-00455486.
PR
XX (UROG-) UROGENESYS INC.
PA
XX A'far DEH, Hubert RS, Raitano AB, Saffran DC, Mitchell SC;
PI Faris M, Jakobovits A;
PI
XX WPI; 2001-367804/38.
DR
XX N-PSDB; AAD07072.
DR

PT New STEAP (six transmembrane epithelial antigen of the prostate)
PT proteins, expressed in human cancers, useful for detecting and treating
PT cancer.
PS
XX Claim 1; Fig 9A-9C; 187pp; English.

CC The present sequence is human six transmembrane epithelial antigen of the
CC prostate (STEAP)-2 protein. STEAP is a member of cell surface serpentine
CC transmembrane antigens. STEAP-2 gene is located on chromosome 7q21 and is
CC used in gene therapy. Inhibiting the development or progression of a
CC cancer (eg. prostate, colon, bladder, lung, ovarian and pancreatic)
CC expressing STEAP or inhibiting growth or killing cells expressing STEAP
CC in a patient, comprises administering a vaccine composition to the
CC patient. Treating a patient with a cancer that expresses STEAP, or
CC inhibiting growth or killing cells expressing STEAP, comprises
CC administering to the patient a vector encoding single chain monoclonal
CC antibody that comprises the variable domains of the heavy and light
CC chains of the monoclonal antibody that specifically binds to STEAP, such
CC that the vector delivers the single chain monoclonal antibody coding
CC sequence to the cancer cells and the encoded single chain monoclonal
CC antibody is expressed intracellularly. Note: This sequence is stated to
CC be the same as that being shown as SEQ ID NO:8 (AAE02841) in sequence
CC listing of the specification. However both the sequences differ at
CC several positions
XX
XX Sequence 454 AA;
SQ

Query Match 100.0%; Score 2351; DB 4; Length 454;
Best Local Similarity 100.0%; Pred. No. 5e-245;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMWGSPPKSLSETCLPENGINGIKDARKVTGVTIGSGDPAKSLTIRLRGCVHYVIGS 60
DB 1 MESISMWGSPPKSLSETCLPENGINGIKDARKVTGVTIGSGDPAKSLTIRLRGCVHYVIGS 60
QY 61 RNPKFASEFPFPHVVDVTHEDALTKNIIFVAIHREHYTSLMDLRHLLVGKILIDVSNM 120
DB 61 RNPKFASEFPFPHVVDVTHEDALTKNIIFVAIHREHYTSLMDLRHLLVGKILIDVSNM 120
QY 121 RINQYBESNAEYIASLFPDGLIVGFNVSAWALQGPXDASRQVYICSNNIQARQVIE 180
DB 121 RINQYBESNAEYIASLFPDGLIVGFNVSAWALQGPXDASRQVYICSNNIQARQVIE 180
QY 181 LARQLNFIPIIDGLISSAREIENPLRLFTLMRGPVVAISLATFFFLYSFVVDVIHPYA 240
DB 181 LARQLNFIPIIDGLISSAREIENPLRLFTLMRGPVVAISLATFFFLYSFVVDVIHPYA 240
QY 241 RNQOSDFYKIPIEIVNKTLPFIVAITLLSLVYLAGLLAAAYQLYGTYKRRFPFMLETWLQ 300
DB 241 RNQOSDFYKIPIEIVNKTLPFIVAITLLSLVYLAGLLAAAYQLYGTYKRRFPFMLETWLQ 300
QY 301 CRKQLGLSFFPAMVAVSYLCLPMRSEERYLFLNMAVQOVHANINENSWNEEVRRIEMV 360
DB 301 CRKQLGLSFFPAMVAVSYLCLPMRSEERYLFLNMAVQOVHANINENSWNEEVRRIEMV 360
QY 361 ISFGIMSLGLSLAIVTSIPSVSNALNWRBESFIQSTLGYVALLISTFHVLIYGWKRAFE 420
DB 361 ISFGIMSLGLSLAIVTSIPSVSNALNWRBESFIQSTLGYVALLISTFHVLIYGWKRAFE 420
QY 421 EBYRRFYTPPNFVALVLPISIVILDLLQLCRYPD 454

DB 421 EBYRRFYTPPNFVALVLPISIVILDLLQLCRYPD 454
RESULT 2
AAU10188
ID AAU10188 standard; protein; 454 AA.
XX
XX AAU10188;
AC
XX
DT 16-JAN-2002 (first entry)
XX
XX Human ORF2 of Six-Transmembrane Protein of Prostate 1.
DE
XX
XX Human; Six-Transmembrane Protein of Prostate 1; STMP1; prostate cancer;
KW benign prostatic hyperplasia; acute prostatitis; testicular cancer;
KW cryptorchidism; testicular disorder; proliferative disorder; lymphoma;
KW leukaemia; melanoma; ovarian cancer; breast cancer; pancreatic cancer;
KW liver cancer; lung cancer; cytostatic; ORF2.
KW
XX
OS Homo sapiens.
OS
XX
XX WO200172962-A2.
PN
XX
XX 04-OCT-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US009410.
PF
XX
XX 24-MAR-2000; 2000US-0191929P.
PR
XX
XX (SAAT/) SAATCIOGLU F.
PA
XX
XX Saetcioglu F;
PI
XX
XX WPI; 2001-662926/76.
DR
XX
XX N-PSDB; AAS15810, AAS15811.
DR
XX

PT New polynucleotide for the diagnosis, prevention and treatment for
PT prostate and testis disorders, particularly prostate cancer, comprises
PT prostate-specific or testis-specific nucleic acids.
XX

XX Claim 1; Fig 4H; 114pp; English.

CC The invention relates to substantially pure prostate-specific or testis-
CC specific polypeptides and the nucleic acids encoding them. Also included
CC are vectors and host cells expressing the proteins, a transgenic animal
CC expressing the protein, antibodies against the proteins, probes for
CC detecting the nucleic acids, antisense molecules for the nucleic acids
CC and methods of isolating modulators of the proteins. Compounds that
CC modulate the prostate specific or testis specific polypeptide are useful
CC to diagnose, prevent or treat disorders of the testis or prostate
CC particularly prostate cancer, benign prostatic hyperplasia, acute
CC prostatitis, testicular cancer, cryptorchidism, undescended, retractile,
CC ascending or vanished testis. Other proliferative disorders for which the
CC modulators may be used include lymphoma, leukaemia, melanoma, ovarian
CC cancer, breast cancer, pancreatic cancer, liver cancer and lung cancer.
CC The present sequence is prostate specific protein, Six-Transmembrane
CC Protein of Prostate 1, STMP1, ORF2
XX
XX Sequence 454 AA;
SQ

Query Match 100.0%; Score 2351; DB 4; Length 454;
Best Local Similarity 100.0%; Pred. No. 5e-245;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMWGSPPKSLSETCLPENGINGIKDARKVTGVTIGSGDPAKSLTIRLRGCVHYVIGS 60
DB 1 MESISMWGSPPKSLSETCLPENGINGIKDARKVTGVTIGSGDPAKSLTIRLRGCVHYVIGS 60
QY 61 RNPKFASEFPFPHVVDVTHEDALTKNIIFVAIHREHYTSLMDLRHLLVGKILIDVSNM 120
DB 61 RNPKFASEFPFPHVVDVTHEDALTKNIIFVAIHREHYTSLMDLRHLLVGKILIDVSNM 120

QY 121 RINQYPESNAEYILASLFPDLSLIVKGFNVSAWALQLGPKDASRQVYICSNNIQARQVIE 180
DB 121 RINQYPESNAEYILASLFPDLSLIVKGFNVSAWALQLGPKDASRQVYICSNNIQARQVIE 180
QY 181 LARQLNFIPIDLGSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
DB 181 LARQLNFIPIDLGSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
QY 241 RNQOSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTXYRRFPFWLETWLQ 300
DB 241 RNQOSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTXYRRFPFWLETWLQ 300
QY 301 CRKQLGLISFFPAMVHAYSLCLPMRSEERYLFNMAYQQOVHANIENSNWNEEVRRIEMY 360
DB 301 CRKQLGLISFFPAMVHAYSLCLPMRSEERYLFNMAYQQOVHANIENSNWNEEVRRIEMY 360
QY 361 ISFGIMSLGLSLLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
DB 361 ISFGIMSLGLSLLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
QY 421 EBYRFRYTPPNFVALVLPISIVILDLQLCRYPD 454
DB 421 EBYRFRYTPPNFVALVLPISIVILDLQLCRYPD 454

RESULT 3

ABG61933
ID ABG61933 standard; protein; 454 AA.
AC ABG61933;

XX 15-AUG-2002 (first entry)

XX Prostate cancer-associated protein #134.

XX Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.

XX Mammalia.

XX WO200230268-A2.

XX 18-APR-2002.

XX 12-OCT-2001; 2001WO-US032045.

XX 13-OCT-2000; 2000US-00687576.

XX 08-DEC-2000; 2000US-00733288.

XX 08-DEC-2000; 2000US-00733742.

XX 24-JAN-2001; 2001US-0263957P.

XX 16-MAR-2001; 2001US-0276791P.

XX 16-MAR-2001; 2001US-0276888P.

XX 06-APR-2001; 2001US-0281922P.

XX 24-APR-2001; 2001US-0286214P.

XX 30-APR-2001; 2001US-00847046.

XX 04-MAY-2001; 2001US-0288589P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;

XX WPI; 2002-471335/50.

XX N-PSDB; ABK92252.

XX Detecting a prostate cancer-associated transcript in a cell in a patient, useful for diagnosing prostate cancer (PC) or screening modulators of PC, by determining if prostate cancer-associated genes are expressed in a prostate tissue.

XX Claim 27; Page 416; 436pp; English.

CC associated polynucleotides (designated PC genes) that selectively hybridise to a sequence that is at least 80% identical to them. The prostate cancer-associated polynucleotide sequences are differentially expressed in prostate tumour tissue or in prostate cancer and are derived from the tissues of various organisms such as humans or other mammals (e.g. mice, sheep and dogs). The methods of the invention are useful for diagnosing and treating prostate cancer in mammals. The prostate cancer-associated genes are useful for diagnosing or treating prostate cancer, as well as for identifying modulators of prostate cancer or agents that inhibit prostate cancer. The nucleic acid sequences are particularly useful in gene therapy, as a vaccine or in antisense applications.

CC ABG61800-ABG61944 represent prostate cancer-associated proteins

CC Sequence 454 AA;

SO Query Match 100.0%; Score 2351; DB 5; Length 454;

Best Local Similarity 100.0%; Pred. No. 5e-245;

Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHWIGS 60

DB 1 MESISMGMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHWIGS 60

QY 61 RNPKEASEFPFHVVDVTHHEDALTNTNIIFVAIHREHYTSLMDLRHLVGLKILIDVSNM 120

DB 61 RNPKEASEFPFHVVDVTHHEDALTNTNIIFVAIHREHYTSLMDLRHLVGLKILIDVSNM 120

QY 121 RINQYPESNAEYILASLFPDLSLIVKGFNVSAWALQLGPKDASRQVYICSNNIQARQVIE 180

DB 121 RINQYPESNAEYILASLFPDLSLIVKGFNVSAWALQLGPKDASRQVYICSNNIQARQVIE 180

QY 181 LARQLNFIPIDLGSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240

DB 181 LARQLNFIPIDLGSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240

QY 241 RNQOSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTXYRRFPFWLETWLQ 300

DB 241 RNQOSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTXYRRFPFWLETWLQ 300

QY 301 CRKQLGLISFFPAMVHAYSLCLPMRSEERYLFNMAYQQOVHANIENSNWNEEVRRIEMY 360

DB 301 CRKQLGLISFFPAMVHAYSLCLPMRSEERYLFNMAYQQOVHANIENSNWNEEVRRIEMY 360

QY 361 ISFGIMSLGLSLLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420

DB 361 ISFGIMSLGLSLLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420

QY 421 EBYRFRYTPPNFVALVLPISIVILDLQLCRYPD 454

DB 421 EBYRFRYTPPNFVALVLPISIVILDLQLCRYPD 454

RESULT 4

AAU76538

ID AAU76538 standard; protein; 454 AA.

XX AAU76538;

XX 05-JUN-2002 (first entry)

XX Tumour-associated antigenic target protein, TAT138.

XX TAT138; Tumour-associated Antigenic Target; tumour; breast cancer;

XX colorectal cancer; lung cancer; ovarian cancer;

XX central nervous system cancer; liver cancer; bladder cancer; melanoma;

XX pancreatic cancer; leukaemia; gene therapy.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Modified-site 8..13 /note= "N-myristoylation site"

XX Modified-site 24..29

FT /note= "N-myristoylation site"
FT Region 29..32
FT /note= "cAMP- and GMP-dependent protein kinase
FT phosphorylation site"
FT Modified-site 34..39
FT /note= "N-myristoylation site"
FT Modified-site 193..198
FT /note= "N-myristoylation site"
FT Domain 210..230
FT /note= "Transmembrane domain"
FT Modified-site 256..259
FT /note= "Asn is N-glycosylated"
FT Domain 257..277
FT /note= "Transmembrane domain"
FT Domain 259..379
FT /note= "Transmembrane domain"
FT Modified-site 274..279
FT /note= "N-myristoylation site"
FT Domain 299..319
FT /note= "Transmembrane domain"
FT Domain 393..413
FT /note= "Transmembrane domain"
FT Modified-site 416..424
FT /note= "Tyrosine kinase phosphorylation site"
FT Domain 428..448
FT /note= "Transmembrane domain"
PN WO200216429-A2.
XX 28-FEB-2002.
XX 22-JUN-2001; 2001WO-US020118.
XX 24-AUG-2000; 2000WO-US023328.
PR 26-SEP-2000; 2000US-0235451P.
PR 01-DEC-2000; 2000WO-US032678.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PA (GETH) GENENTECH INC.
XX Goddard A, Godowski PJ, Gurney AL, Hillan KJ, Polakis P, Smith V;
PI Wood WJ, Wu TD, Zhang Z;
XX WPI; 2002-280917/32.
DR N-PSDB; ABK11093.
XX Novel isolated tumor-associated antigenic target polypeptides which are
PT useful as targets for cancer therapy and diagnosis in mammals.
XX Claim 12; Fig 10; 121pp; English.
XX The invention relates to an isolated tumour-associated antigenic target
CC polypeptide (TAT) (I), specifically TAT134-TAT138 polypeptides, and the
CC polynucleotides (II) encoding them. (II) is useful for diagnosing the
CC presence of a tumour in a mammal, where the level of expression of (II)
CC is indicative on the presence of tumour in the mammal from which the test
CC sample was obtained. Antibody to (I) is useful for killing a cancer cell
CC (e.g. breast cancer cell, a colorectal cancer cell, a lung cancer cell,
CC an ovarian cancer cell, a central nervous system (CNS) cancer cell, a
CC liver cancer cell, a bladder cancer cell, a pancreatic cancer cell, a
CC melanoma cell or a leukaemia cell) that expresses (I). Oligonucleotides
CC hybridising to (II) are useful as diagnostic probes, antisense
CC oligonucleotide probes or for encoding fragments of full length TAT
CC polypeptide. (II) is also useful in chromosome and gene mapping and in
CC the generation of antisense RNA and DNA probes, for constructing
CC hybridisation probes for mapping the gene encoding TAT and for genetic
CC analysis of individuals with genetic disorders. (II) is also useful for
CC generating either transgenic animals or knockout animals, and in gene
CC therapy. The TAT polypeptides and nucleic acids may also be used for
CC tissue typing and the TAT polypeptides are useful for screening compounds
CC that mimic the TAT polypeptide (agonist) or prevent the effect of TAT
CC polypeptide (antagonist). The antibody is useful for staging TAT

CC polypeptide-expressing cancers, purifying or immunoprecipitating TAT
CC polypeptide from cells, for detection and quantitation of TAT polypeptide
CC in vitro, e.g., in an enzyme linked immunosorbent assay (ELISA) or
CC western blot. The antibodies are also useful for treating a TAT-
CC expressing cancer or alleviating one or more symptoms of cancer in a
CC mammal. The present sequence represents the amino acid sequence of TAT138
XX
SQ Sequence 454 AA;
Query Match 100.0%; Score 2351; DB 5; length 454;
Best Local Similarity 100.0%; Pred. No. 5e-245;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MESISWMSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHVIQS 60
DB 1 MESISWMSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHVIQS 60
QY 61 RNPKEAFSEFPFHVDVTHHEDALTKTNIIFVAIHREHTSLMDLRHLVGLIDVSNM 120
DB 61 RNPKEAFSEFPFHVDVTHHEDALTKTNIIFVAIHREHTSLMDLRHLVGLIDVSNM 120
QY 121 RINQYPSNAEYLASLPDLSLYKGFNVSAWALQGPKASROVYICSNIOAQOVIE 180
DB 121 RINQYPSNAEYLASLPDLSLYKGFNVSAWALQGPKASROVYICSNIOAQOVIE 180
QY 181 LARQNFIPIDIGSLSSAREIENLPLRFTLWRGPPVVAISLATFFLYSFVDVHPYA 240
DB 181 LARQNFIPIDIGSLSSAREIENLPLRFTLWRGPPVVAISLATFFLYSFVDVHPYA 240
QY 241 RNQSDFFYKPIEIVNKTLPVATLTLSSLVYLAGLIAAAYQLYYGTYRRPFWLETWQ 300
DB 241 RNQSDFFYKPIEIVNKTLPVATLTLSSLVYLAGLIAAAYQLYYGTYRRPFWLETWQ 300
QY 301 CRKQLGLSFFPAMVHVAYSLCPMRRSERYLFLNMAVQOVHANIEINSWNEBEVRIEM 360
DB 301 CRKQLGLSFFPAMVHVAYSLCPMRRSERYLFLNMAVQOVHANIEINSWNEBEVRIEM 360
QY 361 ISFGIMSLGLSLAVTSIPSVSNALNMRPFSFIQSTLGYVALLISTFHVLIYGWKRAFE 420
DB 361 ISFGIMSLGLSLAVTSIPSVSNALNMRPFSFIQSTLGYVALLISTFHVLIYGWKRAFE 420
QY 421 EYYRFTYPPNFVLAIVLPSTIVLDLQLCRYPD 454
DB 421 EYYRFTYPPNFVLAIVLPSTIVLDLQLCRYPD 454
RESULT 5
AAU80190
ID AAU80190 standard; protein; 454 AA.
XX AAU80190;
XX 15-JUL-2002 (first entry)
XX Human PUMPCn protein, PRO23203.
KW Human; protein upregulated in metastatic prostate cancer; immunogen;
KW PUMPCn; PRO23203; prostate cancer; cytostatic; gene therapy; transgenic;
KW androgen independent prostate cancer; DNA microarray.
OS Homo sapiens.
XX WO200226822-A2.
XX 04-APR-2002.
XX 26-SEP-2001; 2001WO-US030290.
XX 26-SEP-2000; 2000US-0235451P.
XX (GETH) GENENTECH INC.
XX Devaux B, Eberhard D, Goddard A, Godowski PJ, Grimaldi JC;

PI Hillan KJ, Watanabe CK, Wood WI, Yansura DG, Zhang Z;
XX WPI; 2002-383270/41.
DR N-PSDB; ABK50391.
XX
PT New polypeptide termed protein upregulated in metastatic prostate cancer
PT and encoding polynucleotides, useful for identifying polypeptide
PT antagonists for treating prostate cancer.
XX
PS Claim 23; Fig 2; 137pp; English.
XX
CC The invention relates to an isolated human protein upregulated in
CC metastatic prostate cancer (PUMPCn) PRO23203 polypeptide, a sequence 80%
CC identical to PRO23203 and the sequence as encoded by cDNA insert of the
CC vector deposited as ATCC Deposit No. PTA-2513 (DNA185171-2994) on
CC 26/9/2000. Also included are the polynucleotide encoding the protein (or
CC a DNA sequence 80% identical to the polynucleotide and one that
CC hybridises to complement of the polynucleotide), a vector comprising the
CC polynucleotide, a polynucleotide deposited with ATCC under accession
CC number PTA-2513 (DNA185171-2994), a host cell comprising the vector,
CC preparation of PRO23203, a chimeric molecule comprising PRO23203 fused to
CC a heterologous amino acid sequence, an anti-PRO23203 antibody, an
CC agonist/antagonist of PRO23203, and diagnosing the presence of prostate
CC cancer in a mammal by: (a) contacting a microarray diagnostic with a
CC DNA185171-2994 probe, detecting and quantifying hybridisation of
CC DNA185171-2994 probe in prostate cancer tissue compared with normal
CC tissue and determining if DNA185171-2994 is overexpressed; or (b)
CC contacting a tissue of the mammal with an anti-PRO23203 antibody and
CC detecting the binding of the antibody to a component of the tissue, where
CC binding is indicative of the presence of prostate cancer in the mammal.
CC The antibody is useful for treating prostate cancer in mammal which is
CC androgen independent prostate cancer, that has metastasised to another
CC portion of the body, where the antibody is not conjugated with a
CC cytotoxic agent and the method further comprises administering a
CC chemotherapeutic agent to the mammal and for diagnosing the presence of
CC prostate cancer in a mammal. The PRO23203 polynucleotide is useful for
CC generating either transgenic animals or knock out animals which in turn
CC are useful in the development and screening of therapeutically useful
CC reagents or for use in gene therapy and for chromosome identification.
CC PRO23203 and polynucleotide may also be used for tissue typing and
CC PRO23203 may also be employed as a therapeutic agents and for screening
CC compounds to identify those that mimic the PRO23203 polypeptides
CC (agonists) or prevent the effect of the PRO23203 polypeptide
CC (antagonists). PRO23203 is further useful for the affinity purification
CC of PRO23203 from recombinant cell culture or natural sources. The present
CC sequence represents PRO23203
XX
SQ Sequence 454 AA;
Query Match 100.0%; Score 2351; DB 5; Length 454;
Best Local Similarity 100.0%; Pred. No. 5e-245;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MESISMGSPKSLSETCLPENGINGIKDARKVTGVIIGSGDFAKSLTIRLCGYHVVIGS 60
Db 1 MESISMGSPKSLSETCLPENGINGIKDARKVTGVIIGSGDFAKSLTIRLCGYHVVIGS 60
QY 61 RNPKFASEFPFHVVDTTHEDALTNTNIIFVAIHREHYTSLMDLRHLVGKILIDVSNM 120
Db 61 RNPKFASEFPFHVVDTTHEDALTNTNIIFVAIHREHYTSLMDLRHLVGKILIDVSNM 120
QY 121 RINQYPSNAEYIASLFPDSLIVKGFNVSAWALQGPKDASROVYICSNNIQARQVIE 180
Db 121 RINQYPSNAEYIASLFPDSLIVKGFNVSAWALQGPKDASROVYICSNNIQARQVIE 180
QY 181 LARQNLFIPIIDGLSSAREIENLPLRLFTLMRGPVVAISLATFFLYSFVADVHPYA 240
Db 181 LARQNLFIPIIDGLSSAREIENLPLRLFTLMRGPVVAISLATFFLYSFVADVHPYA 240
QY 241 RNOGSDFYKPIPIEIVNKTLPVATLTLVLVLAAGLAAAYQLYYGTYRRFPWLETWLQ 300
Db 241 RNOGSDFYKPIPIEIVNKTLPVATLTLVLVLAAGLAAAYQLYYGTYRRFPWLETWLQ 300

QY 301 CRKQLGLSFFFAWVHVAYSCLPMRRSERYLELNMAVQOVHANIEENNEEVRLEMY 360
Db 301 CRKQLGLSFFFAWVHVAYSCLPMRRSERYLELNMAVQOVHANIEENNEEVRLEMY 360
QY 361 ISFGIMSLGLSLAATVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHYLIGWKRAFE 420
Db 361 ISFGIMSLGLSLAATVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHYLIGWKRAFE 420
QY 421 EEEYRFPYTPPNFVALVLPISIVILDLQLCRYPD 454
Db 421 EEEYRFPYTPPNFVALVLPISIVILDLQLCRYPD 454
RESULT 6
ABU08893
ID ABU08893 standard; protein; 454 AA.
XX AC ABU08893;
XX DT 09-OCT-2003 (first entry)
XX DE Tumour-associated antigenic target 138 (TAT138).
XX KW Human; tumour-associated antigenic target polypeptide; TAT; tumour;
XX gene therapy; cytostatic; tissue typing; prostate tumour; cancer.
XX OS Homo sapiens.
XX PN US2003060612-A1.
XX PD 27-MAR-2003.
XX PF 22-JUN-2001; 2001US-00888257.
XX PR 28-OCT-1997; 97US-0063540P.
XX PR 17-JUN-1998; 98US-0089653P.
XX PR 10-SEP-1998; 98US-0099792P.
XX PR 08-OCT-1998; 98US-0103678P.
XX PR 02-JUN-1999; 99WO-US012252.
XX PR 01-SEP-1999; 99WO-US020111.
XX PR 18-FEB-2000; 2000WO-US004342.
XX PR 02-MAR-2000; 2000WO-US005841.
XX PR 30-MAR-2000; 2000WO-US008439.
XX PR 24-AUG-2000; 2000WO-US023328.
XX PR 26-SEP-2000; 2000US-0235451P.
XX PR 01-DEC-2000; 2000WO-US032678.
XX PR 28-FEB-2001; 2001WO-US006520.
XX PR 01-MAR-2001; 2001WO-US006666.
XX PA (GETH) GEMENTECH INC.
XX PI Goddard A, Godowski PJ, Gurney AL, Hillan KJ, Polakis P, Smith V;
PI Wood WI, Wu TD, Zhang Z;
XX WPI; 2003-576355/54.
XX DR N-PSDB; ACD25893.
XX
PT New nucleic acid, useful for the manufacture of a medicament for
PT diagnosing or treating tumor in a mammal.
XX
PS Claim 12; Fig 10; 71pp; English.
XX
CC The invention discloses human nucleic acids encoding tumour-associated
CC antigenic target (TAT) polypeptides, with or without their associated
CC signal peptide. Also disclosed is an antibody that specifically binds to
CC the TAT polypeptides, a method for detecting the presence of a tumour in
CC a mammal and a method for killing a cancer cell expressing the TAT
CC polypeptide. The nucleotide sequences are useful in preparing TAT
CC polypeptides by recombinant techniques and in gene therapy (e.g. for
CC replacement of defective gene). The TAT polypeptides are useful as
CC therapeutic agents and for detecting the presence, prevention and/or
CC treatment of a tumour, such as colon, breast or prostate tumour. The TAT
CC polypeptides and nucleic acids may also be used diagnostically for tissue

CC typing. The sequence presented is the TAT138 polypeptide of the invention
XX
SQ Sequence 454 AA;

Query Match 100.0%; Score 2351; DB 7; Length 454;
Best Local Similarity 100.0%; Pred. No. 5e-245;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIIGSGDPAKSLTIRLRGCVHYVIGS 60
DB 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIIGSGDPAKSLTIRLRGCVHYVIGS 60
QY 61 RNPKEASEFFPHVVDVTHEDALTKNIIFVAIHREHYTSLMDLRHLLVSKILLIDVSNM 120
DB 61 RNPKEASEFFPHVVDVTHEDALTKNIIFVAIHREHYTSLMDLRHLLVSKILLIDVSNM 120
QY 121 RINQYPSNAEYLALSLFPDSLIVKGFNVSAWALQIGPKDASRQVYICSNNIQARQVIE 180
DB 121 RINQYPSNAEYLALSLFPDSLIVKGFNVSAWALQIGPKDASRQVYICSNNIQARQVIE 180
QY 181 LARQINFIPIIDGLSLSSAREIENLPLRLFTLMRGPVVVAISLATFFFLYSFVRDVIHPYA 240
DB 181 LARQINFIPIIDGLSLSSAREIENLPLRLFTLMRGPVVVAISLATFFFLYSFVRDVIHPYA 240
QY 241 RNOQSDFFYKPIEIVNKTLPVATITLSLVYLAGLAAAYQLYGTKYRRFPFWLETWLQ 300
DB 241 RNOQSDFFYKPIEIVNKTLPVATITLSLVYLAGLAAAYQLYGTKYRRFPFWLETWLQ 300
QY 301 CRKQGLLSFFFAWVAVYSLCLPMRSEERYLFLNMAVQOVHANINENSWNEEVRRIEMY 360
DB 301 CRKQGLLSFFFAWVAVYSLCLPMRSEERYLFLNMAVQOVHANINENSWNEEVRRIEMY 360
QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRRESFIQSTLGYVALLISTFHVLIYGMKRAFE 420
DB 361 ISFGIMSLGLSLAVTSIPSVSNALNWRRESFIQSTLGYVALLISTFHVLIYGMKRAFE 420
QY 421 EBYRFTYTPNPFVLAIVLPSIVILDLQLCRYPD 454
DB 421 EBYRFTYTPNPFVLAIVLPSIVILDLQLCRYPD 454

RESULT 7
ADB65001
ID ADB65001 standard; protein; 454 AA.

AC ADB65001;
XX
DT 04-DEC-2003 (first entry)

DE Human protein encoded by clone PROST20168600.

XX
KW Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;
KW cell regeneration; membrane protein; signal transduction-related protein;
KW transcription-related protein; osteoporosis; neurological disease;
KW cancer; tumor.

XX Homo sapiens.

XX EPI308459-A2.

XX 07-MAY-2003.

XX 28-MAR-2002; 2002EP-00007401.

XX 05-NOV-2001; 2001JP-00379298.

XX 25-JAN-2002; 2002US-00350978.

XX (HELI-) HELIX RES INST.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,
PI Yamamoto J, Isono Y, Otsuka K, Nagai R, Tamechika I,
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y;

XX WPI; 2003-450961/43.
DR N-PSDB; ADB63031.

PT New polynucleotides and polypeptides, useful for developing a diagnostic
PT marker or medicines for regulation of their expression and activity, or
PT as targets of gene therapy.

XX Claim 1; Page; 222pp; English.

CC The invention discloses a polynucleotide comprising a sequence selected
CC from 1970 fully defined nucleotide sequences which encode novel
CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
CC or its partial peptide, an antibody binding to the polypeptide or peptide
CC of the polynucleotide, immunologically assaying the polypeptide or peptide
CC with the antibody of the encoded protein, and observing the binding
CC between the two, a transformant carrying the polynucleotide in an
CC expressible manner and an antisense polynucleotide. The oligonucleotide
CC is useful as a primer for synthesizing the polynucleotide, or as a probe
CC for detecting the polynucleotide. The polynucleotides and encoded
CC proteins are useful as pharmaceutical agents and many disease-related
CC genes may be included in them, for developing a diagnostic marker or
CC medicines for regulation of their expression and activity, or as targets
CC of gene therapy. The genes are involved in tissue and/or cell
CC regeneration. Membrane proteins, signal transduction-related proteins,
CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
CC neurological diseases, cancer, tumors). The cDNA may be used to regulate
CC the activity or expression of the encoded protein to treat diseases. The
CC sequence presented is a protein of the invention. Note: Some of the
CC sequence data for this patent is not represented in the printed
CC specification, but is based on sequence information supplied by the
CC European Patent Office.

XX Sequence 454 AA;

Query Match 99.9%; Score 2348; DB 7; Length 454;
Best Local Similarity 99.8%; Pred. No. 1.1e-244;
Matches 453; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIIGSGDPAKSLTIRLRGCVHYVIGS 60
DB 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIIGSGDPAKSLTIRLRGCVHYVIGS 60
QY 61 RNPKEASEFFPHVVDVTHEDALTKNIIFVAIHREHYTSLMDLRHLLVSKILLIDVSNM 120
DB 61 RNPKEASEFFPHVVDVTHEDALTKNIIFVAIHREHYTSLMDLRHLLVSKILLIDVSNM 120
QY 121 RINQYPSNAEYLALSLFPDSLIVKGFNVSAWALQIGPKDASRQVYICSNNIQARQVIE 180
DB 121 RINQYPSNAEYLALSLFPDSLIVKGFNVSAWALQIGPKDASRQVYICSNNIQARQVIE 180
QY 181 LARQINFIPIIDGLSLSSAREIENLPLRLFTLMRGPVVVAISLATFFFLYSFVRDVIHPYA 240
DB 181 LARQINFIPIIDGLSLSSAREIENLPLRLFTLMRGPVVVAISLATFFFLYSFVRDVIHPYA 240
QY 241 RNOQSDFFYKPIEIVNKTLPVATITLSLVYLAGLAAAYQLYGTKYRRFPFWLETWLQ 300
DB 241 RNOQSDFFYKPIEIVNKTLPVATITLSLVYLAGLAAAYQLYGTKYRRFPFWLETWLQ 300
QY 301 CRKQGLLSFFFAWVAVYSLCLPMRSEERYLFLNMAVQOVHANINENSWNEEVRRIEMY 360
DB 301 CRKQGLLSFFFAWVAVYSLCLPMRSEERYLFLNMAVQOVHANINENSWNEEVRRIEMY 360
QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRRESFIQSTLGYVALLISTFHVLIYGMKRAFE 420
DB 361 ISFGIMSLGLSLAVTSIPSVSNALNWRRESFIQSTLGYVALLISTFHVLIYGMKRAFE 420
QY 421 EBYRFTYTPNPFVLAIVLPSIVILDLQLCRYPD 454
DB 421 EBYRFTYTPNPFVLAIVLPSIVILDLQLCRYPD 454

CC	transmembrane protein differentially expressed in prostate and lung
CC	tumour TMPL
XX	
SQ	Sequence 490 AA;
	Query Match 97.6%; Score 2294; DB 7; Length 490;
	Best Local Similarity 100.0%; Pred. No. 8.4e-239;
	Matches 444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVTIGSGDPAKSLTIRLRGCVHYIGS 60
Db	1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVTIGSGDPAKSLTIRLRGCVHYIGS 60
QY	61 RNPKEFASEFPFHVVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVVKILIDVSNM 120
Db	61 RNPKEFASEFPFHVVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVVKILIDVSNM 120
QY	121 RINQYPESNAEYLASLPDLSLVKGFNVVSAMALQGPKDASROYICSNNIQARQVIE 180
Db	121 RINQYPESNAEYLASLPDLSLVKGFNVVSAMALQGPKDASROYICSNNIQARQVIE 180
QY	181 LARQNFIPIDLGSLSSAREIENLPRLFTLWRGPVVAISLATFFFLYSFVRDVIHPYA 240
Db	181 LARQNFIPIDLGSLSSAREIENLPRLFTLWRGPVVAISLATFFFLYSFVRDVIHPYA 240
QY	241 RNQOSDFYKIPIEIVNKTLPVAILTSLVYLAGLLAAAYQLYGTXYRRFPWLETWLQ 300
Db	241 RNQOSDFYKIPIEIVNKTLPVAILTSLVYLAGLLAAAYQLYGTXYRRFPWLETWLQ 300
QY	301 CRKQGLISFFPAMVHVAYSCLPMRSEERYFLNMAVQOVHANINENSWNEEYVRIEMY 360
Db	301 CRKQGLISFFPAMVHVAYSCLPMRSEERYFLNMAVQOVHANINENSWNEEYVRIEMY 360
QY	361 ISFGIMSLGLSLAVTSIPSVSNALNWRREFSIQSTLGVALLISTFHVLIYGWKRAFE 420
Db	361 ISFGIMSLGLSLAVTSIPSVSNALNWRREFSIQSTLGVALLISTFHVLIYGWKRAFE 420
QY	421 EEYRFFYTPPNFVLALVLPISIVIL 444
Db	421 EEYRFFYTPPNFVLALVLPISIVIL 444
RESULT 10	
AAU10187	
ID	AAU10187 standard; protein; 490 AA.
XX	
AC	AAU10187;
XX	
DT	16-JAN-2002 (first entry)
XX	
DE	Human Six-Transmembrane Protein of Prostate 1, STMPL.
XX	
KW	Human; Six-Transmembrane Protein of Prostate 1; STM; 1; prostate cancer;
KW	benign prostatic hyperplasia; acute prostatitis; testicular cancer;
KW	cryptorchidism; testicular disorder; proliferative disorder; lymphoma;
KW	leukaemia; melanoma; ovarian cancer; breast cancer; pancreatic cancer;
KW	liver cancer; lung cancer; cytostatic.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Domain
FT	Location/Qualifiers
FT	209..230
FT	/label= Transmembrane_domain_1
FT	255..273
FT	/label= Transmembrane_domain_2
FT	304..325
FT	/label= Transmembrane_domain_3
FT	360..380
FT	/label= Transmembrane_domain_4
FT	393..415
FT	/label= Transmembrane_domain_5
FT	432..452
FT	/label= Transmembrane_domain_6

XX	
PN	WO200172962-A2.
XX	
PD	04-OCT-2001.
XX	
PF	23-MAR-2001; 2001WO-US009410.
XX	
PR	24-MAR-2000; 2000US-0191929P.
XX	
PA	(SAAT/) SAATCIOGLU F.
XX	
PI	Saatcioglu F;
XX	
DR	WPI; 2001-662926/76.
DR	N-PSDB; AAS15793, AAS15801, AAS15802.
XX	
PT	New polynucleotide for the diagnosis, prevention and treatment for
PT	prostate and testis disorders, particularly prostate cancer, comprises
PT	prostate-specific or testis-specific nucleic acids.
XX	
PS	Claim 1; Fig 4B; 114pp; English.
XX	
CC	The invention relates to substantially pure prostate-specific or testis-
CC	specific polypeptides and the nucleic acids encoding them. Also included
CC	are vectors and host cells expressing the proteins, a transgenic animal
CC	expressing the protein, antibodies against the proteins, probes for
CC	detecting the nucleic acids, antisense molecules for the nucleic acids
CC	and methods of isolating modulators of the proteins. Compounds that
CC	modulate the prostate specific or testis specific polypeptide are useful
CC	to diagnose, prevent or treat disorders of the testis or prostate
CC	particularly prostate cancer, benign prostatic hyperplasia, acute
CC	prostatitis, testicular cancer, cryptorchidism, undescended, retractile,
CC	ascending or vanished testis. Other proliferative disorders for which the
CC	modulators may be used include lymphoma, leukaemia, melanoma, ovarian
CC	cancer, breast cancer, pancreatic cancer, liver cancer and lung cancer.
CC	The present sequence represents a prostate specific protein, Six-
CC	Transmembrane Protein of Prostate 1, STM;1
XX	
SQ	Sequence 490 AA;
	Query Match 97.4%; Score 2290; DB 4; Length 490;
	Best Local Similarity 99.8%; Pred. No. 2.3e-238;
	Matches 443; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVTIGSGDPAKSLTIRLRGCVHYIGS 60
Db	1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVTIGSGDPAKSLTIRLRGCVHYIGS 60
QY	61 RNPKEFASEFPFHVVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVVKILIDVSNM 120
Db	61 RNPKEFASEFPFHVVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVVKILIDVSNM 120
QY	121 RINQYPESNAEYLASLPDLSLVKGFNVVSAMALQGPKDASROYICSNNIQARQVIE 180
Db	121 RINQYPESNAEYLASLPDLSLVKGFNVVSAMALQGPKDASROYICSNNIQARQVIE 180
QY	181 LARQNFIPIDLGSLSSAREIENLPRLFTLWRGPVVAISLATFFFLYSFVRDVIHPYA 240
Db	181 LARQNFIPIDLGSLSSAREIENLPRLFTLWRGPVVAISLATFFFLYSFVRDVIHPYA 240
QY	241 RNQOSDFYKIPIEIVNKTLPVAILTSLVYLAGLLAAAYQLYGTXYRRFPWLETWLQ 300
Db	241 RNQOSDFYKIPIEIVNKTLPVAILTSLVYLAGLLAAAYQLYGTXYRRFPWLETWLQ 300
QY	301 CRKQGLISFFPAMVHVAYSCLPMRSEERYFLNMAVQOVHANINENSWNEEYVRIEMY 360
Db	301 CRKQGLISFFPAMVHVAYSCLPMRSEERYFLNMAVQOVHANINENSWNEEYVRIEMY 360
QY	361 ISFGIMSLGLSLAVTSIPSVSNALNWRREFSIQSTLGVALLISTFHVLIYGWKRAFE 420
Db	361 ISFGIMSLGLSLAVTSIPSVSNALNWRREFSIQSTLGVALLISTFHVLIYGWKRAFE 420
QY	421 EEYRFFYTPPNFVLALVLPISIVIL 444

Db 421 EEYRYFTPTPNFVLALVLPSTIVL 444

RESULT 11
AAU0189
ID AAU0189 standard; protein; 419 AA.
XX AC AAU0189;
XX DT 16-JAN-2002 (first entry)

XX Human ORF3 of Six-Transmembrane Protein of Prostate 1.
DE Human ORF3 of Six-Transmembrane Protein of Prostate 1.
XX Human; Six-Transmembrane Protein of Prostate 1; STM, 1; prostate cancer;
KW benign prostatic hyperplasia; acute prostatitis; testicular cancer;
KW cryptorchidism; testicular disorder; proliferative disorder; lymphoma;
KW leukaemia; melanoma; ovarian cancer; breast cancer; pancreatic cancer;
KW liver cancer; lung cancer; cytostatic; ORF3.

XX OS Homo sapiens.
XX PN WO200172962-A2.
XX PD 04-OCT-2001.
XX PF 23-MAR-2001; 2001WO-US009410.
XX DT 24-MAR-2000; 2000US-0191929P.
XX PR 24-MAR-2000; 2000US-0191929P.

XX PA (SAAT/) SAATCIOGLU F.
XX PI Saactioglu F;
XX WPI; 2001-662926/76.
XX DR N-PSDB; AAS15813.

XX PT New polynucleotide for the diagnosis, prevention and treatment for
PT prostate and testis disorders, particularly prostate cancer, comprises
PT prostate-specific or testis-specific nucleic acids.

XX PS Claim 1; Fig 4K; 114pp; English.

CC The invention relates to substantially pure prostate-specific or testis-
CC specific polypeptides and the nucleic acids encoding them. Also included
CC are vectors and host cells expressing the proteins, a transgenic animal
CC expressing the protein, antibodies against the proteins, probes for
CC detecting the nucleic acids, antisense molecules for the nucleic acids
CC and methods of isolating modulators of the proteins. Compounds that
CC modulate the prostate specific or testis specific polypeptide are useful
CC to diagnose, prevent or treat disorders of the testis or prostate
CC particularly prostate cancer, benign prostatic hyperplasia, acute
CC prostatitis, testicular cancer, cryptorchidism, undescended, retractile,
CC ascending or vanished testis. Other proliferative disorders for which the
CC modulators may be used include lymphoma, leukaemia, melanoma, ovarian
CC cancer, breast cancer, pancreatic cancer, liver cancer and lung cancer.
CC The present sequence is prostate specific protein, Six-Transmembrane
CC Protein of Prostate 1, STM,1, ORF3

XX SQ Sequence 419 AA;

Query Match 86.6%; Score 2036; DB 4; Length 419;
Best Local Similarity 99.7%; Pred. No. 5.5e-211;
Matches 394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MESISMMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHVG 60
Db 1 MESISMMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHVG 60
OY 61 RNPKPAGSEFPFHVDVTHHEDALTKNIIFVAIHREHYTSLMDLRHLVLGKILLIDVSNM 120
Db 61 RNPKPAGSEFPFHVDVTHHEDALTKNIIFVAIHREHYTSLMDLRHLVLGKILLIDVSNM 120

OY 121 RINQYPSNAEYLAFLPDSLIVKGFNVSAWALQGPDAKROVYICSNNIQARQVIE 180
Db 121 RINQYPSNAEYLAFLPDSLIVKGFNVSAWALQGPDAKROVYICSNNIQARQVIE 180
OY 181 LARQINFPIDGLSLSSAREIENLPLRFTLWRGPVVAISLATEFFLYSFVRDVIHPYA 240
Db 181 LARQINFPIDGLSLSSAREIENLPLRFTLWRGPVVAISLATEFFLYSFVRDVIHPYA 240
OY 241 RNOQSDFFYKPIEIVNKTLPVATLLSLVYLAGLLAAAYQLYYGTXYRRFPFWLETWLQ 300
Db 241 RNOQSDFFYKPIEIVNKTLPVATLLSLVYLAGLLAAAYQLYYGTXYRRFPFWLETWLQ 300
OY 301 CRKQLGLSFFPAMVAVASLCLPWRSEKYLFLNMAQQVYANINENSWNEEVRIMY 360
Db 301 CRKQLGLSFFPAMVAVASLCLPWRSEKYLFLNMAQQVYANINENSWNEEVRIMY 360
OY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRPFSFIQ 395
Db 361 ISFGIMSLGLSLAVTSIPSVSNALNWRPFSFIQ 395

RESULT 12
ABG12306
ID ABG12306 standard; protein; 576 AA.
XX AC ABG12306;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #12297.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX DR N-PSDB; AAS76493.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX PS Claim 20; SEQ ID NO 42665; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 576 AA;

Query Match 83.9%; Score 1972.5; DB 4; Length 576;
Best Local Similarity 87.4%; Pred. No. 6.7e-204;
Matches 390; Conservative 0; Mismatches 1; Indels 55; Gaps 1;

OY 1 MESISMSPKSLSEICLPNGINGIKDARKVTYGVIGSGDFAKSLTIRLRGCVHVG 60
1 MESISMSPKSLSEICLPNGINGIKDARKVTYGVIGSGDFAKSLTIRLRGCVHVG 60
DB 1 MESISMSPKSLSEICLPNGINGIKDARKVTYGVIGSGDFAKSLTIRLRGCVHVG 60
OY 61 RNPFASEFPFHVDTTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
61 RNPFASEFPFHVDTTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
DB 61 RNPFASEFPFHVDTTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
OY 121 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQGPCKASROVYICSNNIQARQVIE 180
121 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQGPCKASROVYICSNNIQARQVIE 180
DB 121 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQGPCKASROVYICSNNIQARQVIE 180
OY 181 LARQNFIPIDLSLSAREIENLPRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
181 LARQNFIPIDLSLSAREIENLPRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
DB 181 LARQNFIPIDLSLSAREIENLPRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
OY 241 RNOQSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTKYRRFPWLETWLQ 300
241 RNOQSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTKYRRFPWLETWLQ 300
DB 241 RNOQSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTKYRRFPWLETWLQ 300
OY 301 CRKQLGLISFFPAMVHVAISLCLPMRSEERYLFNNMAYQOVHANIEENSWNEEVRIMY 360
301 CRKQLGLISFFPAMVHVAISLCLPMRSEERYLFNNMAYQOVHANIEENSWNEEVRIMY 360
DB 301 CRKQLGLISFFPAMVHVAISLCLPMRSEERYLFNNMAYQOVHANIEENSWNEEVRIMY 360
OY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRSEFSIQSTLGYVALLISTFHVLIYGMKRAFE 420
361 ISFGIMSLGLSLAVTSIPSVSNALNWRSEFSIQSTLGYVALLISTFHVLIYGMKRAFE 420
DB 361 ISFGIMSLGLSLAVTSIPSVSNALNWRSEFSIQSTLGYVALLISTFHVLIYGMKRAFE 420
OY 421 EEYRFPYTPPNFVLAIVLPSTIVIDL 446
421 EEYRFPYTPPNFVLAIVLPSTIVIDL 446
DB 421 EEYRFPYTPPNFVLAIVLPSTIVIDL 446
OY 366 EEYRFPYTPPNFVLAIVLPSTIVIDL 391
366 EEYRFPYTPPNFVLAIVLPSTIVIDL 391
DB 366 EEYRFPYTPPNFVLAIVLPSTIVIDL 391

RESULT 13
ABG00113
ID ABG00113 standard; protein; 1273 AA.
XX
AC ABG00113;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #104.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX

PI Drmanac RT, Liu C, Tang YT;
XX
XX WPI: 2001-639362/73.
DR N-PSDB; AAS64300.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 30472; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1273 AA;

Query Match 83.9%; Score 1972.5; DB 4; Length 1273;
Best Local Similarity 87.4%; Pred. No. 2.3e-203;
Matches 390; Conservative 0; Mismatches 1; Indels 55; Gaps 1;

OY 1 MESISMSPKSLSEICLPNGINGIKDARKVTYGVIGSGDFAKSLTIRLRGCVHVG 60
1 MESISMSPKSLSEICLPNGINGIKDARKVTYGVIGSGDFAKSLTIRLRGCVHVG 60
DB 1 MESISMSPKSLSEICLPNGINGIKDARKVTYGVIGSGDFAKSLTIRLRGCVHVG 60
OY 73 MESISMSPKSLSEICLPNGINGIKDARKVTYGVIGSGDFAKSLTIRLRGCVHVG 132
73 MESISMSPKSLSEICLPNGINGIKDARKVTYGVIGSGDFAKSLTIRLRGCVHVG 132
DB 73 MESISMSPKSLSEICLPNGINGIKDARKVTYGVIGSGDFAKSLTIRLRGCVHVG 132
OY 61 RNPFASEFPFHVDTTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
61 RNPFASEFPFHVDTTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
DB 61 RNPFASEFPFHVDTTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
OY 133 RNPFASEFPFHVDTTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 192
133 RNPFASEFPFHVDTTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 192
DB 133 RNPFASEFPFHVDTTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 192
OY 121 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQGPCKASROVYICSNNIQARQVIE 180
121 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQGPCKASROVYICSNNIQARQVIE 180
DB 121 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQGPCKASROVYICSNNIQARQVIE 180
OY 193 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQGPCKASROVYICSNNIQARQVIE 252
193 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQGPCKASROVYICSNNIQARQVIE 252
DB 193 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQGPCKASROVYICSNNIQARQVIE 252
OY 181 LARQNFIPIDLSLSAREIENLPRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
181 LARQNFIPIDLSLSAREIENLPRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
DB 181 LARQNFIPIDLSLSAREIENLPRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
OY 253 LARQNFIPIDLSLSAREIENLPRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 312
253 LARQNFIPIDLSLSAREIENLPRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 312
DB 253 LARQNFIPIDLSLSAREIENLPRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 312
OY 241 RNOQSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTKYRRFPWLETWLQ 300
241 RNOQSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTKYRRFPWLETWLQ 300
DB 241 RNOQSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTKYRRFPWLETWLQ 300
OY 313 RNOQSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTKYRRFPWLETWLQ 372
313 RNOQSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTKYRRFPWLETWLQ 372
DB 313 RNOQSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTKYRRFPWLETWLQ 372
OY 301 CRKQLGLISFFPAMVHVAISLCLPMRSEERYLFNNMAYQOVHANIEENSWNEEVRIMY 360
301 CRKQLGLISFFPAMVHVAISLCLPMRSEERYLFNNMAYQOVHANIEENSWNEEVRIMY 360
DB 301 CRKQLGLISFFPAMVHVAISLCLPMRSEERYLFNNMAYQOVHANIEENSWNEEVRIMY 360
OY 373 CRKQLGLISFFPAMVHVAISLCLPMRSEERYLFNNMAYQOVHANIEENSWNEEVRIMY 411
373 CRKQLGLISFFPAMVHVAISLCLPMRSEERYLFNNMAYQOVHANIEENSWNEEVRIMY 411
DB 373 CRKQLGLISFFPAMVHVAISLCLPMRSEERYLFNNMAYQOVHANIEENSWNEEVRIMY 411
OY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRSEFSIQSTLGYVALLISTFHVLIYGMKRAFE 420
361 ISFGIMSLGLSLAVTSIPSVSNALNWRSEFSIQSTLGYVALLISTFHVLIYGMKRAFE 420
DB 361 ISFGIMSLGLSLAVTSIPSVSNALNWRSEFSIQSTLGYVALLISTFHVLIYGMKRAFE 420
OY 412 -----QSTLGYVALLISTFHVLIYGMKRAFE 437
412 -----QSTLGYVALLISTFHVLIYGMKRAFE 437
DB 412 -----QSTLGYVALLISTFHVLIYGMKRAFE 437
OY 421 EEYRFPYTPPNFVLAIVLPSTIVIDL 446
421 EEYRFPYTPPNFVLAIVLPSTIVIDL 446
DB 421 EEYRFPYTPPNFVLAIVLPSTIVIDL 446
OY 438 EEYRFPYTPPNFVLAIVLPSTIVIDL 463
438 EEYRFPYTPPNFVLAIVLPSTIVIDL 463
DB 438 EEYRFPYTPPNFVLAIVLPSTIVIDL 463

RESULT 14
AAE02841
ID AAE02841 standard; protein; 450 AA.
XX
AC AAE02841;
XX
DT 06-AUG-2001 (first entry)
XX
DE Human STEAP-2 protein, alternative version.
XX
KW Human; cytostatic; antiproliferative; vaccine; gene therapy;
XX six transmembrane epithelial antigen of the prostate-2; STEAP-2;
KW chromosome 7q21; cancer; prostate; colon; bladder; lung; ovarian;
KW pancreatic.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 335..336
FT /note= "Encoded by GACTGAGCT"
FT Misc-difference 375..376
FT /note= "Encoded by AGATGAAGT"
FT Misc-difference 415..416
FT /note= "Encoded by GAGTAAAGC"
FT Misc-difference 445..446
FT /note= "Encoded by ACATGAAGT"
FT Misc-difference 448..449
FT /note= "Encoded by AATTAAATTC"
XX
PN WO200140276-A2.
XX
PD 07-JUN-2001.
XX
PF 06-DEC-2000; 2000WO-US033040.
XX
PR 06-DEC-1999; 99US-00455486.
XX
PA (UROG-) UROGENESYS INC.
XX
PI Afar DEH, Hubert RS, Raitano AB, Safiran DC, Mitchell SC;
PI Faris M, Jakobovits A;
XX
DR WPI; 2001-367804/38.
DR N-PSDB; AAD07072.
XX
PT New STEAP (six transmembrane epithelial antigen of the prostate)
PT proteins, expressed in human cancers, useful for detecting and treating
PT cancer.
XX
PS Claim 1; Page 175-176; 187pp; English.
XX
CC The present sequence is an alternative version of human six transmembrane
CC epithelial antigen of the prostate (STEAP)-2 protein. STEAP is a member
CC of cell surface serpentine transmembrane antigens. STEAP-2 gene is
CC located on chromosome 7q21 and is used in gene therapy. Inhibiting the
CC development or progression of a cancer (eg. prostate, colon, bladder,
CC lung, ovarian and pancreatic) expressing STEAP or inhibiting growth or
CC killing cells expressing STEAP in a patient, comprises administering a
CC vaccine composition to the patient. Treating a patient with a cancer that
CC expresses STEAP, or inhibiting growth or killing cells expressing STEAP,
CC comprises administering to the patient a vector encoding single chain
CC monoclonal antibody that comprises the variable domains of the heavy and
CC light chains of the monoclonal antibody that specifically binds to STEAP,
CC such that the vector delivers the single chain monoclonal antibody coding
CC sequence to the cancer cells and the encoded single chain monoclonal
CC antibody is expressed intracellularly. Note: This sequence is stated to
CC be the same as that being shown as SEQ ID NO:8 (AAE02781) in figure 9A-9C
CC of the specification. However the present sequence lacks several amino
CC acids at its N-terminal end and has additional amino acids at its C-
CC terminal end when compared with the sequence shown in figure 9A-9C
XX
SQ Sequence 450 AA;

Query Match 73.9%; Score 1738; DB 4; Length 450;
Best Local Similarity 100.0%; Pred. No. 1.1e-178;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 119 NMRINQVPSNAEYLASLPDGLIVGFNVSAWALQGPDKASRQVYICSNNIQARQV 178
DB 1 NMRINQVPSNAEYLASLPDGLIVGFNVSAWALQGPDKASRQVYICSNNIQARQV 60
QY 179 TELARQINFIPIIDGLSSAREIENPLRFTLMRGPVVAISLATFFLYSFVRDVIHP 238
DB 61 TELARQINFIPIIDGLSSAREIENPLRFTLMRGPVVAISLATFFLYSFVRDVIHP 120
QY 239 YARNQSDFFYKIPIEIVNKTLPVATITLISLVYLAGLLAAVQLYGTYKRRFPWLETW 298
DB 121 YARNQSDFFYKIPIEIVNKTLPVATITLISLVYLAGLLAAVQLYGTYKRRFPWLETW 180
QY 299 LQCRKQILGSLFFPAMVHVAYSLCLPWRSEKYLFLNMAVQOVHANIENSWNEEVRWIE 358
DB 181 LQCRKQILGSLFFPAMVHVAYSLCLPWRSEKYLFLNMAVQOVHANIENSWNEEVRWIE 240
QY 359 MYISFGIMSLGLSLLAVTSIPSVSNALNWRFSFIQSTIGYVALLISTFHVLIYGWKRA 418
DB 241 MYISFGIMSLGLSLLAVTSIPSVSNALNWRFSFIQSTIGYVALLISTFHVLIYGWKRA 300
QY 419 FEEERYRFTYPPNFVLAIVLPISIVLDLQLCRYPD 454
DB 301 FEEERYRFTYPPNFVLAIVLPISIVLDLQLCRYPD 336
RESULT 15
ABP62883
ID ABP62883 standard; protein; 1082 AA.
XX
AC ABP62883;
XX
DT 14-OCT-2002 (first entry)
XX
DE Human polypeptide SEQ ID NO 320.
XX
KW Human; vulnery; dermatological; neuroprotective; nootropic; cancer;
KW antiparkinsonian; immunostimulant; cytostatic; immunosuppressive;
KW antidiabetic; antiallergic; gene therapy; wound healing; tissue repair;
KW burn; central nervous system disorder; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease; immune disorder;
KW autoimmune disorder; multiple sclerosis; diabetes; allergy.
XX
OS Homo sapiens.
XX
PN WO200218424-A2.
XX
PD 07-MAR-2002.
XX
PF 31-AUG-2001; 2001WO-US027093.
XX
PR 01-SEP-2000; 2000US-00654935.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Asundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J;
PI Zhao QA, Wang D, Liu C, Drmanac RT, Wehrman T;
XX
DR WPI; 2002-583321/62.
DR N-PSDB; ABQ93362.
XX
PT New polynucleotide and polypeptides, useful for treatment and diagnosis
PT of Alzheimer's, Parkinson's, Huntington's, amyotrophic lateral
PT sclerosis, immune deficiencies, cancer, autoimmune disorders, multiple
PT sclerosis, diabetes and allergies.
XX
PS Claim 20; SEQ ID NO 320; 284pp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide (I) comprising one of
CC 245 sequences (ABQ93288-ABQ93532). Treating a condition comprising

CC administering to a mammalian subject a composition comprising the protein
 CC (II) encoded by (I) (ABP62809-ABP63053) or an antibody (III) to (II).
 CC (I), (II) and (III) are useful for diagnostic evaluation of disorders.
 CC (I) is useful for gene therapy of diseases and (II) can be used for
 CC therapeutic treatment. Diseases that may be treated include wound healing
 CC and tissue repair, burns, central nervous system disorders (e.g.
 CC Alzheimer's, Parkinson's, Huntington's and amyotrophic lateral
 CC sclerosis), immune deficiencies, cancer, autoimmune disorders, multiple
 CC sclerosis, diabetes and allergies. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

Sequence 1082 AA;

Query Match	58.0%;	Score 1364.5;	DB 5;	Length 1082;
Best Local Similarity	83.2%;	Pred. No. 1,1e-137;		
Matches 272;	Conservative	0;	Mismatches 0;	Indels 55;
				Gaps 1;

QY	120	MRINQYBESNAEYLLASLEFPDLSLVKGFNVVSAMALQLGPKDASRQVYICSNNIQARQOVI	179
Db	1	MRINQYBESNAEYLLASLEFPDLSLVKGFNVVSAMALQLGPKDASRQVYICSNNIQARQOVI	60
QY	180	ELARQINFIPIIDLGSLSSAREIENLPRLFTLMRGFPVVAISLATFFELYSFVRDVIHPY	239
Db	61	ELARQINFIPIIDLGSLSSAREIENLPRLFTLMRGFPVVAISLATFFELYSFVRDVIHPY	120
QY	240	ARNQSDPYKIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYGTXRFRPEPMLETWL	299
Db	121	ARNQSDPYKIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYGTXRFRPEPMLETWL	180
QY	300	QCRKQJGLLSFFEFAMVHVAYSJCLPMRPRSERYJFLNMAYYQOVHANTENSWNEEVRITEM	359
Db	181	QCRKQJGLLSFFEFAMVHVAYSJCLPMRPRSERYJFLNMAYYQ-----	220
QY	360	YISFGIMSLGLLSLLAVTSPSVSNALNWREFSFOSTLGVYALLISTFHVLIYGMKRAF	419
Db	221	-----OSTLGVYALLISTFHVLIYGMKRAF	245
QY	420	EEBYRFRFYTPNPFVLLAVLPSIVILDL	446
Db	246	EEBYRFRFYTPNPFVLLAVLPSIVILDL	272

Search completed: April 16, 2004, 12:39:11
Job time : 63 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 12:36:32 ; Search time 46 Seconds
(without alignments)
3114.027 Million cell updates/sec

Title: US-09-455-486-6
Perfect score: 2351
Sequence: 1 MESISMGSPKSLSETCLPN.....ALVLPISIVLIDLQLCRYPD 454

Scoring table: BUCSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPREMBL_25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2290	97.4	490	4	Q8IUE7
2	2284	97.2	490	4	Q8NFT2
3	2239.5	95.3	489	11	Q8BWB6
4	1318	56.1	488	11	Q8C5F0
5	1316	56.0	488	11	Q99P41
6	1316	56.0	488	11	Q8CIS9
7	1316	56.0	526	11	Q80ZF3
8	1306	55.6	514	11	Q924Z1
9	1272	54.1	488	4	Q86SF6
10	1272	54.1	498	4	Q7Z389
11	1269	54.0	498	4	Q9NVB5
12	1252.5	53.3	487	4	Q8NEW6
13	1245.5	53.0	487	4	Q8TF03
14	1074.5	45.7	456	4	Q8TDP3
15	1064.5	45.3	470	11	Q923B6
16	1057.5	45.0	470	11	Q91W31

ID	Q8IUE7	PRELIMINARY	PRT	490 AA.	Q91ZE8	Q91ZE8 mus musculu
17	1041.5	44.3	474	11	Q91ZE8	Q91ZE8 mus musculu
18	956.5	40.7	464	11	Q7TP88	Q7TP88 rattus norv
19	728	31.0	338	6	Q9GL50	Q9GL50 sus scrofa
20	715	30.4	339	11	Q924Z2	Q924Z2 mus musculu
21	707	30.1	339	11	Q9CWR7	Q9CWR7 mus musculu
22	704	29.9	339	11	Q924J9	Q924J9 mus musculu
23	628	26.7	283	4	Q9H5R1	Q9H5R1 mus sapien
24	541	23.0	283	4	Q8WNB0	Q8WNB0 homo sapien
25	324	13.8	143	4	Q9H7Y1	Q9H7Y1 homo sapien
26	171.5	7.3	208	16	Q8NLU6	Q8NLU6 corynebacte
27	170	7.2	208	16	Q8XQS2	Q8XQS2 ralsconia s
28	164.5	7.0	234	16	Q82AX0	Q82AX0 streptomyc
29	160	6.8	198	16	Q930K7	Q930K7 rhizobium m
30	159.5	6.8	211	16	Q8YK44	Q8YK44 anabaena sp
31	156.5	6.7	239	16	Q9RKR6	Q9RKR6 streptomyc
32	154.5	6.6	225	2	Q8GFG4	Q8GFG4 rhodococcus
33	151	6.4	200	16	Q98C62	Q98C62 rhizobium 1
34	149.5	6.4	218	2	Q8GFG3	Q8GFG3 rhodococcus
35	149.5	6.4	221	16	Q82Q14	Q82Q14 streptomyc
36	147.5	6.3	226	2	Q8GHR6	Q8GHR6 rhodococcus
37	147	6.3	213	17	Q29059	Q29059 archaeoglob
38	144.5	6.1	226	2	Q9AH05	Q9AH05 rhodococcus
39	143.5	6.1	223	17	Q8TYQ9	Q8TYQ9 methanopyru
40	142	6.0	212	17	Q29370	Q29370 archaeoglob
41	142	6.0	224	1	Q59661	Q59661 methanobact
42	137.5	5.8	199	2	Q9F418	Q9F418 mycobacteri
43	136.5	5.8	217	2	Q83VCL	Q83VCL streptomyc
44	136.5	5.8	217	16	Q9JN78	Q9JN78 streptomyc
45	135.5	5.8	203	2	Q8GFG1	Q8GFG1 rhodococcus

ALIGNMENTS

RESULT 1
Q8IUE7
ID Q8IUE7
AC Q8IUE7
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE STAMPl.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RX MEDLINE=22229309; PubMed=12095985;
RA Korkmaz K.S., Elbi C.C., Korkmaz C.G., Loda M., Hager G.L.,
RA Saaticioglu F.;
RT "Molecular cloning and characterization of STAMPl, a highly prostate
RT specific six-trans-membrane protein that is overexpressed in prostate
RT cancer.";
RL J. Biol. Chem. 277:36689-36696(2002).
DR EMBL; AY008445; AAC32149.1; -
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420_oxidored; I.
SQ SEQUENCE 490 AA; 56027 MW; D1971A84880F7E51 CRC64;

Query Match 97.4%; Score 2290; DB 4; Length 490;
Best Local Similarity 99.8%; Pred. No. 41e-176;
Matches 443; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	MESISMGSPKSLSETCLPN	INGIKDARKTVGVIGSGDFAKSLTIRLCGYHVIGS	60
QY	1	MESISMGSPKSLSETCLPN	INGIKDARKTVGVIGSGDFAKSLTIRLCGYHVIGS	60
Db	1	MESISMGSPKSLSETCLPN	INGIKDARKTVGVIGSGDFAKSLTIRLCGYHVIGS	60
QY	61	RNPKFASEFP	PHVVDVTHHEDALTKNTIIFVAIHREHYTSLWDLRHLVGLIDVSNM	120
Db	61	RNPKFASEFP	PHVVDVTHHEDALTKNTIIFVAIHREHYTSLWDLRHLVGLIDVSNM	120


```

QY 121 RINQYPSNAEYLAFLPDSLIVKGFNVSAWALQGPKDASRQVYICSNNIQARQVIE 180
    |||
Db 121 RINQYPSNAEYLAFLPDSLIVKGFNVSAWALQGPKDASRQVYICSNNIQARQVIE 180
QY 181 LARQNFIPIDGLSLSSAREIENLPLRLFTLMRGVVAISLATFFLYSFVRDVIHPYA 240
    |||
Db 181 LARQNFIPIDGLSLSSAREIENLPLRLFTLMRGVVAISLATFFLYSFVRDVIHPYA 240
QY 241 RNOQSDFFYKPIEIVNKTLPVIAITLLSLVYLAGLLAAAYQLYGTYKRRFPFWLETWLQ 300
    |||
Db 241 RNOQSDFFYKPIEIVNKTLPVIAITLLSLVYLAGLLAAAYQLYGTYKRRFPFWLETWLQ 300
QY 301 CRKQGLSLFFFAVHVAYSLCLPMRSEERYLFLNMAYYQVHANINIENSWNEEVRILEMY 360
    |||
Db 301 CRKQGLSLFFFAVHVAYSLCLPMRSEERYLFLNMAYYQVHANINIENSWNEEVRILEMY 360
QY 361 ISFGIMSLGLSLSLAVTSIPSVSNALNWRREFSFIOSTLGYVALLISTFHVLIYGMKRAFE 420
    |||
Db 361 ISFGIMSLGLSLSLAVTSIPSVSNALNWRREFSFIOSTLGYVALLISTFHVLIYGMKRAFE 420
QY 421 EBYRFFYTPPNFVLALVLPISIVIL 444
    |||
Db 421 EBYRFFYTPPNFVLALVLPISIVIL 444

```

RESULT 2

```

Q8NFT2 PRELIMINARY; PRT; 490 AA.
ID Q8NFT2;
AC Q8NFT2;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Six-transmembrane epithelial antigen of prostate 2.
GN STEAP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RX [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Portka K.P., Helenius M.A., Visakorpi T.;
RT "Cloning and characterization of a novel six-transmembrane protein
RT STEAP2, down-regulated in androgen-independent prostate cancer.";
RT Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF455138; AAN04080.1; -.
DR GeneW; HGNC:17885; STEAP2.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420_oxidored; 1.
KW Transmembrane.
SQ SEQUENCE 490 AA; 55961 MW; F2E9C30CDACCEF81 CRC64;

```

Query Match 97.2%; Score 2284; DB 4; Length 490;
 Best Local Similarity 99.8%; Pred. No. 1.2e-175;
 Matches 443; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHVG 60
    |||
Db 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHVG 60
QY 61 RNPKFASEFPFHVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
    |||
Db 61 RNPKFASEFPFHVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
QY 121 RINQYPSNAEYLAFLPDSLIVKGFNVSAWALQGPKDASRQVYICSNNIQARQVIE 180
    |||
Db 121 RINQYPSNAEYLAFLPDSLIVKGFNVSAWALQGPKDASRQVYICSNNIQARQVIE 180
QY 181 LARQNFIPIDGLSLSSAREIENLPLRLFTLMRGVVAISLATFFLYSFVRDVIHPYA 240
    |||

```

```

Db 181 LARQNFIPIDGLSLSSAREIENLPLRLFTLMRGVVAISLATFFLYSFVRDVIHPYA 240
QY 241 RNOQSDFFYKPIEIVNKTLPVIAITLLSLVYLAGLLAAAYQLYGTYKRRFPFWLETWLQ 300
    |||
Db 241 RNOQSDFFYKPIEIVNKTLPVIAITLLSLVYLAGLLAAAYQLYGTYKRRFPFWLETWLQ 300
QY 301 CRKQGLSLFFFAVHVAYSLCLPMRSEERYLFLNMAYYQVHANINIENSWNEEVRILEMY 360
    |||
Db 301 CRKQGLSLFFFAVHVAYSLCLPMRSEERYLFLNMAYYQVHANINIENSWNEEVRILEMY 360
QY 361 ISFGIMSLGLSLSLAVTSIPSVSNALNWRREFSFIOSTLGYVALLISTFHVLIYGMKRAFE 420
    |||
Db 361 ISFGIMSLGLSLSLAVTSIPSVSNALNWRREFSFIOSTLGYVALLISTFHVLIYGMKRAFE 420
QY 421 EBYRFFYTPPNFVLALVLPISIVIL 444
    |||
Db 421 EBYRFFYTPPNFVLALVLPISIVIL 444

```

RESULT 3

```

Q8WB6 PRELIMINARY; PRT; 489 AA.
ID Q8WB6;
AC Q8WB6;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE weakly similar to tumor suppressor PHYE.
GN 4921538B17RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK052981; BAC35230.1; -.
DR MGD; MGI:1921301; 4921538B17RIK.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420_oxidored; 1.
SQ SEQUENCE 489 AA; 55760 MW; 98CD63D59DDDF24C CRC64;

```

Query Match 95.3%; Score 2239.5; DB 11; Length 489;
 Best Local Similarity 97.3%; Pred. No. 4.8e-172;
 Matches 432; Conservative 9; Mismatches 2; Indels 1; Gaps 1;

```

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHVG 60
    |||
Db 1 MESISMGSPKSL-ETFLPNGINGIKDARQVTGVIGSGDFAKSLTIRLRGCVHVG 59
QY 61 RNPKFASEFPFHVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
    |||
Db 61 RNPKFASEFPFHVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 119
QY 121 RINQYPSNAEYLAFLPDSLIVKGFNVSAWALQGPKDASRQVYICSNNIQARQVIE 180
    |||
Db 121 RINQYPSNAEYLAFLPDSLIVKGFNVSAWALQGPKDASRQVYICSNNIQARQVIE 179
QY 181 LARQNFIPIDGLSLSSAREIENLPLRLFTLMRGVVAISLATFFLYSFVRDVIHPYA 240
    |||
Db 181 LARQNFIPVDGLSSAKAEIENLPLRLFTLMRGVVAISLATFFLYSFVRDVIHPYA 239
QY 241 RNOQSDFFYKPIEIVNKTLPVIAITLLSLVYLAGLLAAAYQLYGTYKRRFPFWLETWLQ 300
    |||
Db 241 RNOQSDFFYKPIEIVNKTLPVIAITLLSLVYLAGLLAAAYQLYGTYKRRFPFWLETWLQ 299
QY 301 CRKQGLSLFFFAVHVAYSLCLPMRSEERYLFLNMAYYQVHANINIENSWNEEVRILEMY 360
    |||

```

```
DB 300 CRKQGLISFFFAVHVAYSLCLPMRSEKYLFLNMAYQVHANENAMNEEYVRIEMY 359
QY 361 ISFGINSLGLLSLAVTSIPSVSNALNWRBESFIQSTLGYVALLISTFHVLIYGMKRAFE 420
DB 360 ISFGINSLGLLSLAVTSIPSVSNALNWRBESFIQSTLGYVALLITTFHVLIYGMKRAFA 419
QY 421 EBYRYTPPNFVLALVLPISIVIL 444
DB 420 EBYRYTPPNFVLALVLPISIVIL 443

RESULT 4
Q8C5F0 PRELIMINARY; PRT; 488 AA.
ID Q8C5F0
AC Q8C5F0;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE DUDULIN 2 homolog.
GN TSAP6 OR 101001D01RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK078769; BAC37383.1; -.
DR MGD; MGI:1915678; Tsap6.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420 oxidored; 1.
SQ SEQUENCE 488 AA; 54780 MW; FDFC4F42AE503D11 CRC64;

Query Match 56.1%; Score 1318; DB 11; Length 488;
Best Local Similarity 56.2%; Pred. No. 8.4e-98;
Matches 248; Conservative 81; Mismatches 100; Indels 12; Gaps 4;

QY 6 MMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVIGSRNPKF 65
DB 14 LVDSDGSLAE--VP-----KEAPK--VGILGSGDFARSLATRLVSGSFVVVGSRNPKR 63
QY 66 ASEFPFHVVDTVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVGKILIDVSNMRIN-- 123
DB 64 TAGLFPSLAQVTFQEEAMSSPEVIFVAVRHHYSSLSLADQLAGKILVDVSNPTEKERL 123
QY 124 QYPESNAEYTLASLFPDLSLVKGFNVSAVALQGPKDASROVYICSNNIQARQVIELAR 183
DB 124 QHRQSNAEYTLASLFPACTVVKAFNVIISAWALQAGPRDGNROVLICSDQPEAKRTISEMAR 183
QY 184 QLNFIPIDLGSLSSAREIENLPLRLFTLWGRPVVAISLATFFFLYSFVRDVIHPYARNQ 243
DB 184 AMGFTPLDMGSLASAREVEAIPRLPLPSWKVPTLLALGLFVCFYTYNFIKRDVLPYIRKD 243
QY 244 QSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTGYRRPFPWLETWLQCRK 303
DB 244 ENKFFYKMPLSVNTTLPVAVYLLSLVYLPVLAALQLRRGTGYRRPFDWLDMWLQHRK 303
QY 304 QLGILSFFFAVHVAYSLCLPMRSEKYLFLNMAYQVHANENAMNEEYVRIEMYISF 363
DB 304 QIGILSFFFAVHVAYSLCLPLRRSHRYDLVNLAVKQVLANKSRMLWEEVWRMEIYLSL 363
QY 364 GIMSGLSLAVALTSIPSVSNALNWRBESFIQSTLGYVALLISTFHVLIYGMKRAFEERY 423
DB 364 GVLALGMLSLAVALTSIPSIANSINLWKEFSFVQSTLGFVALILSTMHTLTYGWTTRAFEENH 423
```

```
QY 424 YRFYTPPNFVLALVLPISIVIL 444
DB 424 YKFYLPPTFTLLLPVCIIL 444

RESULT 5
Q99P41 PRELIMINARY; PRT; 488 AA.
ID Q99P41
AC Q99P41;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Tumor suppressor pHyde.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Dunning; TISSUE=Prostatic carcinoma;
RX MEDLINE=20424188; PubMed=10969787;
RA Steiner M.S., Zhang X., Wang Y., Lu Y.;
RT "Growth inhibition of prostate cancer by an adenovirus expressing a
RT novel tumor suppressor gene, pHyde.";
RL Cancer Res. 60:4419-4425(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Dunning; TISSUE=Prostatic carcinoma;
RA Lu Y., Rinaldy A.R., Steiner M.S.;
RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Steiner M.S., Rinaldy A.;
RT "A novel prostate cancer associated gene.";
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF335281; AAX00361.1; -.
DR EMBL; AF238865; AAL78207.1; -.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420 oxidored; 1.
SQ SEQUENCE 488 AA; 54640 MW; 59FF07121919FDCB CRC64;

Query Match 56.0%; Score 1316; DB 11; Length 488;
Best Local Similarity 56.2%; Pred. No. 1.2e-97;
Matches 248; Conservative 81; Mismatches 100; Indels 12; Gaps 4;

QY 6 MMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVIGSRNPKF 65
DB 14 LVDSDGSLAE--VP-----KEAPK--VGILGSGDFARSLATRLVSGSFVVVGSRNPKR 63
QY 66 ASEFPFHVVDTVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVGKILIDVSNMRIN-- 123
DB 64 TAGLFPSLAQVTFQEEAVSSPEVIFVAVRHHYSSLSLADQLAGKILVDVSNPTEKERL 123
QY 124 QYPESNAEYTLASLFPDLSLVKGFNVSAVALQGPKDASROVYICSNNIQARQVIELAR 183
DB 124 QHRQSNAEYTLASLFPACTVVKAFNVIISAWALQAGPRDGNROVLICSDQPEAKHTVSEMAR 183
QY 184 QLNFIPIDLGSLSSAREIENLPLRLFTLWGRPVVAISLATFFFLYSFVRDVIHPYARNQ 243
DB 184 AMGFTPLDMGSLASAREVEAIPRLPLPSWKVPTLLALGLSTQSYAANFIKRDVLPYIRKD 243
QY 244 QSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTGYRRPFPWLETWLQCRK 303
DB 244 ENKFFYKMPLSVNTTLPVAVYLLSLVYLPVLAALQLRRGTGYRRPFDWLDMWLQHRK 303
QY 304 QLGILSFFFAVHVAYSLCLPMRSEKYLFLNMAYQVHANENAMNEEYVRIEMYISF 363
DB 304 QIGILSFFFAVHVAYSLCLPLRRSHRYDLVNLAVKQVLANKSRMLWEEVWRMEIYLSL 363
QY 364 GIMSGLSLAVALTSIPSVSNALNWRBESFIQSTLGYVALLISTFHVLIYGMKRAFEERY 423
DB 364 GVLALGMLSLAVALTSIPSIANSINLWKEFSFVQSTLGFVALMLSTMHTLTYGWTTRAFEENH 423
```


QY 424 YRFYTPNFFVLAIVLPSIVIL 444
DB 424 YKFYLPPTFTLTLLPCVIL 444

RESULT 6

Q8C159 PRELIMINARY; PRT; 488 AA.
AC Q8C159;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Similar to RIKEN cDNA 1010001D01 gene.
GN TSAP6 OR 1010001D01RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CZECH II;
RA Strausberg R.;
RL Submitted (SEP-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; BC037435; AAH37435.1; -.
DR MGD; MGI:1915678; Tsap6.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420 oxidored; 1.
SQ SEQUENCE 488 AA; 54749 MW; 9A08D99C90CF83F4 CRC64;

Query Match 56.0%; Score 1316; DB 11; Length 488;
Best Local Similarity 56.2%; Pred. No. 1.2e-97;
Matches 248; Conservative 81; Mismatches 100; Indels 12; Gaps 4;

QY 6 MMGSPKSLSETCPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVIGSRNPKF 65
DB 14 LVDSDSLAE--VP-----KEAPK--VGILGSGDFARSLATRLVSGSFVVVGSRNPKR 63
QY 66 ASEFFPHVVDVTHEDALTKTNIIFVAIHREHYTSLWDLRHLVSKILLIDVSNMRIN-- 123
DB 64 TAGLFPSLAQVTFQEEAVSSPEVIFVAVREHYSSLSCLADQLAGKILLVDVSNPTEKEHL 123
QY 124 QYPESNAEYLASLFPDLSLVKGFNVVSAMALQGPKASROVYICSNNIQARQVIELAR 183
DB 124 QHRQSNAEYLASLFPACTVKAFFNVISAMALQAGPRDGNROVLICSDQPEAKRTISEMAR 183
QY 184 QLNFIPIDLGSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYARNQ 243
DB 184 AMGFTPLDMGSLASAREVEAIPRLPLPSWKVFTLLALGLFVCFYTYNFIRDVLQPIRKD 243
QY 244 QSDFYKIPIEIVNKLPIVAITLLSLVYLAGLAAAYQLYGTGYRRPPELWTLOCRK 303
DB 244 ENKFYKMPLSVNTTILPCVAYVLLSLVYLPGLAALQLRRGTYQRPDMDLHWLOHRK 303
QY 304 QIGLLSFFFAFAMTHALYSFCLPLRRSHRYDLVNLAVKQVLANKSRLWVEEVRMEIYLSL 363
DB 304 QIGLLSFFFAFAMTHALYSFCLPLRRSHRYDLVNLAVKQVLANKSRLWVEEVRMEIYLSL 363
QY 364 GIMSGLSLAVALTSPSVSNALNWRFEFSIQSTLGYVALLISTFHVLIYGWKRAFEERY 423
DB 364 GVALGMLSLAVALTSPSIANSINWKEFSVQSTLGFVALILSTMHTLTYGWTRAFEEENH 423
QY 424 YRFYTPNFFVLAIVLPSIVIL 444
DB 424 YKFYLPPTFTLTLLPCVIL 444

RESULT 7
Q80ZF3 PRELIMINARY; PRT; 526 AA.
AC Q80ZF3;
DT 01-JUN-2003 (TReMBLrel. 24, Created)

DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Tsap6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSMZ ACC331; TISSUE=Bone marrow;
RX MEDLINE=22506415; PubMed=12606722;
RA Passer B.J., Nancy-Portebois V., Amzallag N., Prieur S., Cans C.,
RA Roborel de Climens A., Ficuci G., Bouvard V., Tynnder M., Susini L.,
RA Morchoisne S.P., Cribie V., Lespagnol A., Dausset J., Oren M.,
RA Amson R., Telerman A.;
RT "The p53-inducible TSAP6 gene product regulates apoptosis and the cell
RT cycle and interacts with Nix and the Myt1 kinase."
RL Proc. Natl. Acad. Sci. U.S.A. 100:2284-2289(2003).
DR EMBL; AY214462; AA038239.1; -.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420 oxidored; 1.
SQ SEQUENCE 526 AA; 58530 MW; 6306CD717E25200A CRC64;

Query Match 56.0%; Score 1316; DB 11; Length 526;
Best Local Similarity 56.2%; Pred. No. 1.3e-97;
Matches 248; Conservative 81; Mismatches 100; Indels 12; Gaps 4;

QY 6 MMGSPKSLSETCPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVIGSRNPKF 65
DB 52 LVDSDSLAE--VP-----KEAPK--VGILGSGDFARSLATRLVSGSFVVVGSRNPKR 101
QY 66 ASEFFPHVVDVTHEDALTKTNIIFVAIHREHYTSLWDLRHLVSKILLIDVSNMRIN-- 123
DB 102 TAGLFPSLAQVTFQEEAVSSPEVIFVAVREHYSSLSCLADQLAGKILLVDVSNPTEKEHL 161
QY 124 QYPESNAEYLASLFPDLSLVKGFNVVSAMALQGPKASROVYICSNNIQARQVIELAR 183
DB 162 QHRQSNAEYLASLFPACTVKAFFNVISAMALQAGPRDGNROVLICSDQPEAKRTISEMAR 221
QY 184 QLNFIPIDLGSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYARNQ 243
DB 222 AMGFTPLDMGSLASAREVEAIPRLPLPSWKVFTLLALGLFVCFYTYNFIRDVLQPIRKD 281
QY 244 QSDFYKIPIEIVNKLPIVAITLLSLVYLAGLAAAYQLYGTGYRRPPELWTLOCRK 303
DB 282 ENKFYKMPLSVNTTILPCVAYVLLSLVYLPGLAALQLRRGTYQRPDMDLHWLOHRK 341
QY 304 QIGLLSFFFAFAMTHALYSFCLPLRRSHRYDLVNLAVKQVLANKSRLWVEEVRMEIYLSL 363
DB 342 QIGLLSFFFAFAMTHALYSFCLPLRRSHRYDLVNLAVKQVLANKSRLWVEEVRMEIYLSL 401
QY 364 GIMSGLSLAVALTSPSVSNALNWRFEFSIQSTLGYVALLISTFHVLIYGWKRAFEERY 423
DB 402 GVALGMLSLAVALTSPSIANSINWKEFSVQSTLGFVALILSTMHTLTYGWTRAFEEENH 461
QY 424 YRFYTPNFFVLAIVLPSIVIL 444
DB 462 YKFYLPPTFTLTLLPCVIL 482

RESULT 8
Q924Z1 PRELIMINARY; PRT; 514 AA.
AC Q924Z1;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Dudulin 2.
GN TSAP6 OR 1010001D01RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

[illegible]

```

RESULT 11
Q9NVB5
ID Q9NVB5 PRELIMINARY; PRT; 488 AA.
AC Q9NVB5;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Hypothetical protein FLJ10829 (Dudulin 2).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Serru V., Manivet P., Lenoir C., Eschwege P., Lamblin D.,
RA Vabourdoille M., Kellermann O., Loric S.;
RT "Dudulin 2, a new tumor antigen expressed in various human tumors.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK001691; BAA91839.1; -.
DR EMBL; AY029585; AAK50538.1; -.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420_oxidored; I.
DR PROSITE; PS00290; Ig_MHC; 1.
KM Hypothetical protein.
SQ SEQUENCE 488 AA; 54616 MW; BCOBCA483335AAD6 CRC64;

```

```

Query Match      54.0%; Score 1269; DB 4; Length 488;
Best Local Similarity 54.3%; Pred. No. 7.5e-94;
Matches 241; Conservative 83; Mismatches 108; Indels 12; Gaps 4;

QY      3 SISMGSPKSLSETCLPNGINGIKDARKVTGVIIGSGDFAKSLTIRLIRCGYHVGSRN 62
      11 SLHLVDSDSLAK--VPD-----EAPK--VGILGSGDFARSLATRLVSGGFKVVGSRN 60
QY      63 PKFASEFFPHVVDVTHHEDALTKNTIIFVAIHREHYTSLMDLRHLVLGKILIDVSN--NM 120
      61 PKRTARLYPSAQVTFQEEAVSSPEVIFVAIFREHYSSLCSLSDLQLAGKILVDVSNPTQ 120
Db      121 RINQYPPESNAEYTLASLPFDSLIVKGFNVSAMALQLGPKDASROYICSNNIQARQVIE 180
      121 EHLQHRESNAEYTLASLPFCTVKAFNVISAWTLQAGPRDGNRQVPICGDQPEAKRAVSE 180
QY      181 LARQLNFIPIDLGSISSAREIENLPLRLFTLWRGPVVAISLATFFFLYSFVRDVIHPYA 240
      181 MALAMGFMPVDMGSLASAMEVEAMPLRLPAWKVPTLLALGLFVCFYAYNFVRDVLQPYV 240
Db      241 RNQOSDFYKPIPIEIVNKTLPPIVAITLLSLVYLAGLAAAYQLYYGTYRRFPFWLETWLQ 300
      241 QESQNKFFKLPVSVVNTTLLPCVAYVLLSLVYLPGLVLAALQLRGTKYQRFPDWLDHWLQ 300
QY      301 CRKQLGLLSFFFAVTVAVASLCLPMRRESRYLFLNMAYQOVHANIEINSWNEEVRRIEMY 360
      301 HRKQIGLLSFFCAALHALYSFCLPLRAHRYDLVNLAVKQVLANKSHLWVEEVRWEIY 360
QY      361 ISFGIMSLGLSLAVTISPSVSANLNMWREFSFQSTLCGVALLISTFHVLIYGWKRAFE 420
      361 LSLGLVALGTLISLAVTSLPSIANSLNMWREFSFVQSLGVALVSLTHTLTLYGWTRAFE 420
Db      421 EYYRFTYTPPNFVLAIVLPSIVIL 444
      421 ESRYKFTYLPFTFTLLVPCVIL 444

```

```

RESULT 12
Q8NEW6
ID Q8NEW6 PRELIMINARY; PRT; 487 AA.
AC Q8NEW6;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, last annotation update)
DE PHYde.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lu Y., Beheshti B., Squire J.A., Yang X.J.;
RT "Characterization of a novel apoptosis-inducing gene, hphYde, that
RT inhibits prostate cancer cell growth.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY082673; AAM45136.1; -.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420_oxidored; I.
DR PROSITE; PS00290; IG_MHC; 1.
SQ SEQUENCE 487 AA; 54501 MW; E0EF95E855C81EDF CRC64;

Query Match          53.3%; Score 1252.5; DB 4; Length 487;
Best Local Similarity 54.1%; Pred. No. 1.6e-92;
Matches 240; Conservative 83; Mismatches 108; Indels 13; Gaps 5

QY      3 SISMMGSPKSLSETCLPENGINGIKDARKVTGVIGSGDFAKSLTIRLRICGYHVVIGSRN 62
   |: :| | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db     11 SLHLVSDSSLAK--VPD-----EAPKVS--ILGSGDFARSLATRLVGSGFKVWVGSRN 60
        63 PKFASEFFPHVVDVTHHEDALTKTNIIFVAIHREHYTSLWDLRHLVGKILLDVSN--NM 120

```


Db 61 PKRTARLFPSAAQVTFQEEAVSSPEVIFVAVFREHYSSLCSDQLAGKILVDVSNPTQ 120

QY 121 RINQYPSNAEYLAFLPDSLIVKGNVVSAMALQGPKASROYICSNNOARQOYIE 180

Db 121 EHLQHRRESNAEYLAFLPCTCTVKAFNVISAWTLQAGPRDGNQVPICGDQPEAKRAVSE 180

QY 181 LARQNLFPIDGLSSAREIENLPLRLFTLMRGPVVAISLATFFLYSFVRDVIHPYA 240

Db 181 MALAMGFMPVDMGSLASAMEVEAMPRLPLPAWKVPTLLALGLFVCFYAYNFVRDVLQPYV 240

QY 241 RNQOSDFYKIPBIEIVNKTLPVATLLSLVYLAGLLAAAYQLYYGTXYRFPFWLETWLQ 300

Db 241 QESQNKFFKLPFVSVNNTLLPCVAYVLLSLVYLPGLAAALQLRGTYQRFPDWLDHWLQ 300

QY 301 CRKQGLLSFFPAMVHAYSLCLPMRSEERYLFNMAYQOVHANIENSWNEEVRRIEMY 360

Db 301 HRKQIGLISFFCAALHALYSFCLPLRAHRYDLVNLAVKQVLANKSHLM-VEEVRMEIY 359

QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRSEFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420

Db 360 LSLGVALGTLISLAVTSLPSIANSINWREBSFVQSSLGVALVLTHTLTYGWTBFAFE 419

QY 421 EBYRFTYTPNFVLAALVPSIVIL 444

Db 420 ESRYKFYLPPTFTLLVPCVIL 443

RESULT 13

Q8TF03 PRELIMINARY; PRT; 487 AA.

AC Q8TF03; 01-JUN-2002 (TREMBlrel. 21, Created)

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

DE Tumor suppressor pHyde.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Steiner M.S., Allay J.A., Wang C.;

RT "A novel prostate-derived tumor suppressor.";

RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF238864; AAL78206.1; "

DR GO; GO:0006118; P:electron transport; IEA.

DR InterPro; IPR003006; IG_MHC.

DR InterPro; IPR004455; NADPoxred_F420.

DR Pfam; PF03807; F420_oxidored; I.

DR PROSITE; PS00290; IG_MHC; 1.

SQ SEQUENCE 487 AA; 54402 MW; AFF16053590E6F68 CRC64;

Query Match 53.0%; Score 1245.5; DB 4; Length 487;

Best Local Similarity 53.8%; Pred. No. 5.8e-92;

Matches 239; Conservative 83; Mismatches 109; Indels 13; Gaps 5;

QY 3 SISMGSPKSLSETCLPNGINGIKDARKYTVGVIGSGDPAKSLTIRLIRCGYHVIGSRN 62

Db 11 SLHLVSDSSILAK--VPD-----EAPKVS--ILGSGDFARSLATRLVSGFKVVGSRN 60

QY 63 PKFASEFFPHVVDVTHEDALTNTNIFVAIHREHYTSLMDLRHLVGLIDVSN--NM 120

Db 61 PKRTARLFPSAAQVTFQEEAVSSPEVIFVAVFREHYSSLCSDQLAGKILVDVSNPTQ 120

QY 121 RINQYPSNAEYLAFLPDSLIVKGNVVSAMALQGPKASROYICSNNOARQOYIE 180

Db 121 EHLQHRRESNAEYLAFLPCTCTVKAFNVISAWTLQAGPRDGNQVPICGDQPEAKRAVSE 180

QY 181 LARQNLFPIDGLSSAREIENLPLRLFTLMRGPVVAISLATFFLYSFVRDVIHPYA 240

Db 181 MALAMGFMPVDMGSLASAMEVEAMPRLPLPAWKVPTLLALGLFVCFYAYNFVRDVLQPYV 240

QY 241 RNQOSDFYKIPBIEIVNKTLPVATLLSLVYLAGLLAAAYQLYYGTXYRFPFWLETWLQ 300

Db 241 QESQNKFFKLPFVSVNNTLLPCVAYVLLSLVYLPGLAAALQLRGTYQRFPDWLDHWLQ 300

QY 301 CRKQGLLSFFPAMVHAYSLCLPMRSEERYLFNMAYQOVHANIENSWNEEVRRIEMY 360

Db 301 HRKQIGLISFFCAALHALYSFCLPLRAHRYDLVNLAVKQVLANKSHLM-VEEVRMEIY 359

QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRSEFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420

Db 360 LSLGVALGTLISLAVTSLPSIANSINWREBSFVQSSLGVALVLTHTLTYGWTBFAFE 419

QY 421 EBYRFTYTPNFVLAALVPSIVIL 444

Db 420 ESRYKFYLPPTFTLLVPCVIL 443

RESULT 14

Q8TDP3 PRELIMINARY; PRT; 456 AA.

AC Q8TDP3; 01-JUN-2002 (TREMBlrel. 21, Created)

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

DE PHYDE II.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Wang C., Allay J.A., Steiner M.S.;

RT "Second human member of pHyde family, Human pHyde II.";

RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF262322; AAM08128.1; "

DR GO; GO:0006118; P:electron transport; IEA.

DR InterPro; IPR003006; IG_MHC.

DR InterPro; IPR004455; NADPoxred_F420.

DR Pfam; PF03807; F420_oxidored; I.

DR PROSITE; PS00290; IG_MHC; 1.

SQ SEQUENCE 456 AA; 50430 MW; C5F7C7008D55251E CRC64;

Query Match 45.7%; Score 1074.5; DB 4; Length 456;

Best Local Similarity 52.9%; Pred. No. 3.2e-78;

Matches 209; Conservative 75; Mismatches 98; Indels 13; Gaps 5;

QY 3 SISMGSPKSLSETCLPNGINGIKDARKYTVGVIGSGDPAKSLTIRLIRCGYHVIGSRN 62

Db 11 SLHLVSDSSILAK--VPD-----EAPKVS--ILGSGDFARSLATRLVSGFKVVGSRN 60

QY 63 PKFASEFFPHVVDVTHEDALTNTNIFVAIHREHYTSLMDLRHLVGLIDVSN--NM 120

Db 61 PKRTARLFPSAAQVTFQEEAVSSPEVIFVAVFREHYSSLCSDQLAGKILVDVSNPTQ 120

QY 121 RINQYPSNAEYLAFLPDSLIVKGNVVSAMALQGPKASROYICSNNOARQOYIE 180

Db 121 EHLQHRRESNAEYLAFLPCTCTVKAFNVISAWTLQAGPRDGNQVPICGDQPEAKRAVSE 180

QY 181 LARQNLFPIDGLSSAREIENLPLRLFTLMRGPVVAISLATFFLYSFVRDVIHPYA 240

Db 181 MALAMGFMPVDMGSLASAMEVEAMPRLPLPAWKVPTLLALGLFVCFYAYNFVRDVLQPYV 240

QY 241 RNQOSDFYKIPBIEIVNKTLPVATLLSLVYLAGLLAAAYQLYYGTXYRFPFWLETWLQ 300

Db 241 QESQNKFFKLPFVSVNNTLLPCVAYVLLSLVYLPGLAAALQLRGTYQRFPDWLDHWLQ 300

QY 301 CRKQGLLSFFPAMVHAYSLCLPMRSEERYLFNMAYQOVHANIENSWNEEVRRIEMY 360

Db 301 HRKQIGLISFFCAALHALYSFCLPLRAHRYDLVNLAVKQVLANKSHLM-VEEVRMEIY 359

QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRSEFSFIQ 395

Db 360 LSLGVALGTLISLAVTSLPSIANSINWREBSFVQ 394

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 12:33:12 ; Search time 17 Seconds

(without alignments)
1390.579 Million cell updates/sec

Title: US-09-455-486-6

Perfect score: 2351

Sequence: 1 MESISMGMGSPKSLSETCLPN.....ALVLPSTIVILDLLQLCRYPD 454

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	717	30.5	339	1 STEA_HUMAN	Q9uhe8 homo sapien
2	133	5.7	223	1 F4RE_MERUA	Q58896 methanococ
3	120.5	5.1	232	1 F4RE_MERTH	Q26350 methanobact
4	111.5	4.7	695	1 FSHR_MACTA	P32212 macaca fasc
5	109.5	4.7	206	1 Y538_PAGMU	Q9cn97 pasteurella
6	109	4.6	693	1 FSHR_CHICK	P79763 gallus gall
7	109	4.6	712	1 FRE6_YEAST	Q12473 saccharomyc
8	107.5	4.6	629	1 FRE7_YEAST	Q12333 saccharomyc
9	107.5	4.6	695	1 FSHR_HUMAN	P23945 homo sapien
10	104.5	4.4	695	1 FSHR_PIG	P49059 sus scrofa
11	104	4.4	396	1 SOTB_ECO57	P58529 escherichia
12	104	4.4	396	1 SOTB_ECOLI	P31122 escherichia
13	103	4.4	465	1 NPT1_MOUSE	Q61983 mus musculu
14	102.5	4.4	391	1 NUCC_NEPOL	Q9tkv6 nephroscelm
15	102	4.3	220	1 Y304_BRUMB	Q8ydv3 bruceella me
16	102	4.3	220	1 Y3J1_BRUST	Q8fv59 bruceella su
17	100	4.3	476	1 RF3_SACBA	P05512 saccharomyc
18	99.5	4.2	452	1 NU4M_BRALA	Q79421 branchiosto
19	99	4.2	365	1 SPR1_HUMAN	Q15743 homo sapien
20	98	4.2	3411	1 POLG_YEFLV	P03314 y genome po
21	98	4.2	3411	1 POLG_YEFLV	P19901 y genome po
22	97.5	4.1	268	1 YD49_AQUAE	P35364 aquifex ae
23	97.5	4.1	695	1 FSHR_BOVIN	P35376 bos taurus
24	97	4.1	218	1 YG46_XANAC	Q8p1y9 xanthomonas
25	97	4.1	692	1 FSHR_RAT	P20395 rattus norv
26	96.5	4.1	686	1 FRE1_YEAST	P32791 saccharomyc
27	96	4.1	261	1 PHSC_ECOLI	P77409 escherichia
28	95	4.0	202	1 YAU1_PSEAB	Q9hvas pseudomonas
29	95	4.0	393	1 NUCC_ARATH	P56753 arabidopsis
30	95	4.0	393	1 NUCC_SPIOL	Q9m315 spinacia ol
31	95	4.0	692	1 FSHR_MOUSE	P35378 mus musculu
32	94.5	4.0	492	1 SECY_CYAPA	P25014 cyanophora
33	94.5	4.0	695	1 FSHR_SHEEP	P35379 ovis aries

34	93.5	4.0	365	1 SPR1_MOUSE	Q8bfq3 mus musculu
35	93.5	4.0	452	1 NU4M_BRALF	Q47423 branchiosto
36	93.5	4.0	499	1 GSHR_PLAF7	Q15770 plasmodium
37	93.5	4.0	530	1 AAAL_MOUSE	Q9jmh8 mus musculu
38	93.5	4.0	1233	1 VLI_REOVD	P15024 reovirus (t
39	93	4.0	388	1 HMC3_DESVH	P33390 desulfovibr
40	93	4.0	459	1 NU4M_BOVIN	P03910 bos taurus
41	93	4.0	521	1 YT25_CABEL	Q10934 caenorhabdi
42	92.5	3.9	490	1 GTRI_CHICK	P46896 gallus gall
43	92.5	3.9	500	1 NU4C_OENHO	P58419 oenothera h
44	92.5	3.9	883	1 YHL6_YEAST	P38781 saccharomyc
45	92.5	3.9	1769	1 YUK9_YEAST	P42945 saccharomyc

ALIGNMENTS

RESULT 1	STEA_HUMAN	STANDARD;	PRT;	339 AA.
ID	STEA_HUMAN			
AC	Q9UHE8; Q95034;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Six transmembrane epithelial antigen of prostate.			
GN	STEAP OR STEAP1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20056277; PubMed=10588738;			
RA	Hubert R.S., Vivanco I., Chen E., Rastegar S., Leong K.,			
RA	Mitchell S.C., Madraswala R., Zhou Y., Kuo J., Raitano A.B.,			
RA	Jakobovits A., Saffran D.C., Afar D.E.H.;			
RT	"STEAP: a prostate-specific cell-surface antigen highly expressed in			
RT	human prostate tumors.";			
RI	Proc. Natl. Acad. Sci. U.S.A. 96:14523-14528(1999).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Abu-Threideh J., Stoneking T., Langston Y., Maupin R.;			
RI	Submitted (JUN-1998) to the EMBL/Genbank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Strausberg R.L., Feilgold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Urdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Bosak S.A., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahney J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,			
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,			
RA	Schneerch A., Schein J.E., Jones S.J.M., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length			
RT	human and mouse cDNA sequences.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).			
CC	-1- TISSUE SPECIFICITY: Highly expressed in prostate tumors.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AF186249; AAF17479.1; -
DR EMBL; AC005053; AAC79150.1; ALT_INIT.
DR EMBL; AC004969; AAD15620.2; -
DR EMBL; BC011802; AAH11802.1; -
DR Genew; HGNC:11378; STEAD.
DR MIM; 604415; -
DR GO; GO:0005887; C:integral to plasma membrane; TAS.
DR GO; GO:0005911; C:intercellular junction; TAS.
DR GO; GO:0015267; F:channel/pore class transporter activity; TAS.
KW Transmembrane; Antigen.
FT TRANSMEM 71 91 POTENTIAL.
FT TRANSMEM 119 139 POTENTIAL.
FT TRANSMEM 164 184 POTENTIAL.
FT TRANSMEM 218 238 POTENTIAL.
FT TRANSMEM 258 278 POTENTIAL.
FT TRANSMEM 291 311 POTENTIAL.
SQ SEQUENCE 339 AA; 39851 MW; 55443A170C870387 CRC64;

Query Match 30.5%; Score 717; DB 1; Length 339;
Best Local Similarity 54.9%; Pred. No. 6e-48; Indels 0; Gaps 0;
Matches 130; Conservative 48; Mismatches 59;

QY 208 LFTLMRGPVVVAISLATEFFELYSFVRDVTHPYARNQSDFFYKIPPIVNTLPIVAITLL 267
DB 67 LFPQWHLPIKIAITIASLTFLYTLREVTHPLATSHQYFYKIPPIVNTLPIVNTL 126
QY 268 SLVYLAGLAAAYQLYYGTXYRFPFWLEFWLQCRKQLGLSFFFAWVHVAYSCLPMR 327
DB 127 ALVYLPGVIAIVQLHNGTKYKFPFWLDMKWLTRKQGLSFFFAVHAIVSLSYPMR 186
QY 328 SERXYLFINMAYQOVHANIEENSWNEEVRLEMYISFGIMSLGLSLAVTSIPSVSNALN 387
DB 187 SYRYKLINMAYQOVQNKEDAWIEHDVWRMEIVYSLGIVGLAIALAVTSIPSVSDSLT 246
QY 388 WREFSFIQSTLGYVALLISTFHVLIYGVKRAFESEYRYFTPPNFVALVLPISVIL 444
DB 247 WREFHYIQSKLGIYVLLGTIHALIFAWNKWIDIKQFVWYTPPTFNLAVFLPIVILI 303

RESULT 2

F4RE_METJA STANDARD; PRT; 223 AA.
ID F4RE_METJA
AC Q58836;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative F420-dependent NADP reductase (EC 1.-.-.-).
GN MJ1501.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073 (1996).
CC -!- FUNCTION: Catalyzes the reduction of F420 with NADP(+) and the

CC reduction of NADP(+) with F420H(2). Probably couples the NADP-
CC dependent oxidation of the alcohol to the aldehyde with the F420-
CC dependent reduction of CO(2) to methane (anabolic function) (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: NADPH + coenzyme F420 = NADP(+) + reduced
CC coenzyme F420.
CC -!- SIMILARITY: TO M.ORGANOPHILUM F420-DEPENDENT NADP REDUCTASE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U67591; AAB99514.1; -
DR PIR; D64487; D64487.
DR TIGR; MJ1501; -
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420 oxidored; 1.
DR TIGRFAMs; TIGR00301; TIGR00301; 1.
KW Hypothetical protein; Oxidoreductase; NADP; Complete proteome.
SQ SEQUENCE 223 AA; 24068 MW; 2370BBD6F5BBD3D9 CRC64;

Query Match 5.7%; Score 133; DB 1; Length 223;
Best Local Similarity 21.9%; Pred. No. 0.003;
Matches 46; Conservative 52; Mismatches 78; Indels 34; Gaps 7;

QY 37 GSGDFAKSLTIRLCGYHVVIGSRNPKASEFPFHVVDV-----THEDALT 84
DB 8 GTGDQGFGLARLAK-NNKIIIGSRKKEAEPAKAKEILKQGIADIIIGLENKDAK 66
QY 85 KNTIIFVAIHREH-YTSLWDLRHLVGLIIVD-----SNNRINQPE-SNAEYL 133
DB 67 EGDVILSLPYEYTLSTIKQLKEELKGIIVSIGVPLATAGDKPRLLPFDGVAEMV 126
QY 134 ASLEPDSLIVKGVNVSAWALQLGPKDASROYIICSNINQARQVIELARQLNFI-PIDL 192
DB 127 QNVLESKVSAPQNVCHAVLEDLDPVDCDILVCGNDEAKKVYIDLANQIDGVRAIDC 186
QY 193 GSLSAREIENLPLRLFTLMRGPVVVAISL 222
DB 187 GNLEKSRRIIEAIT-----PLIIGLNI 207

RESULT 3

F4RE_METTH STANDARD; PRT; 232 AA.
ID F4RE_METTH
AC O26350;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative F420-dependent NADP reductase (EC 1.-.-.-).
GN MTH248.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzaden R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RA "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155 (1997).

CC -1- FUNCTION: Catalyzes the reduction of F420 with NADP(+) and the
CC reduction of NADP(+) with F420H(2). Probably couples the NADP-
CC dependent oxidation of the alcohol to the aldehyde with the F420-
CC dependent reduction of CO(2) to methane (anabolic function) (By
CC similarity).

CC -1- CATALYTIC ACTIVITY: NADPH + coenzyme F420 = NADP(+) + reduced
CC coenzyme F420.

CC -1- SIMILARITY: TO M.ORGANOPHILUM F420-DEPENDENT NADP REDUCTASE.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AE000811; AAB84754.1; -.
DR PIR; A69131; A69131.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420_oxidored; 1.
DR TIGRFAMs; TIGR00301; TIGR00301; 1.
KW Hypothetical protein; Oxidoreductase; NADP; Complete proteome.
SQ SEQUENCE 232 AA; 24539 MW; A1CE60ABC8474296 CRC64;

Query Match	5.1%;	Score 120.5;	DB 1;	Length 232;
Best Local Similarity	23.8%;	Pred. No. 0.029;		
Matches 49;	Conservative 42;	Mismatches 90;	Indels 25;	Gaps 6;

QY		30	KATVGIVIGSGDEPAKSLTIRLIRCQGYHVVIGSRNPKFASFEFFPHVDVTHE-----	80
Db		8	KIAV-IGGTGDOGIGIALRFAVAGEEVIIGSRDAEKASKAASKYLEIAGRDDISVECATN	66
QY		81	-DALTKNTIIFAAIH-REHYTSLMDRLHLVGKILIDVS-----NNMRINQYPESN	1299
Db		67	PDAASADVVVLTVPLQAQMWTLASIRDQVRDKVLIDATVPIDSCIGGSAVRYIDWEGS	126
QY		130	AEYLASLF---PDSLIVKGENVVSAMAWLOLPKODASRQVYTCSNNIQARQOVIELARQLN	186
Db		127	AAERAARFLREQGTRVAALAFNNNISASALLEVSEPVDCDLVASPDHRDALLEVAELAEKID	186
QY		187	FI-PIDLGLSSAREIENLPLRLFLL	211
Db		187	GVARIECGGLENAARIIEKITPLLINL	212

RESULT 4

FSHR	MACFA	STANDARD;	PRT;
ID	FSHR	MACFA	695 AA.

DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Follicle stimulating hormone receptor precursor (FSH-R) (Folliotropin
 DE receptor).
 GN FSHR.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=94071854; PubMed=7504463;
 RA Gremoll J., Dankbar B., Sharma R.S., Nieschlag E.;
 RT "Molecular cloning of the testicular follicle stimulating hormone
 RT receptor of the non human primate Macaca fascicularis and
 RT identification of multiple transcripts in the testis.";
 RL Biochem. Biophys. Res. Commun. 196;1066-1072(1993).
 CC -1- FUNCTION: Receptor for follicle stimulating hormone. The activity
 CC of this receptor is mediated by G proteins which activate
 CC adenylyate cyclase.

```

CC      -|- SUBCELLULAR LOCATION: Integral membrane protein.
CC      -|- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC      FSH/LSH/TSH subfamily.
CC      -|- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).

```

DR EMBL; X74454; CAA52463.1; -.
DR PIR; JN0898; JN0898.
DR HSSP; P23945; 1XUN.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR_3.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PRO0237; GPCR RHODOPSIN.
DR SMART; SM00013; LRRNT; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Phosphorylation; Repeat; Leucine-rich repeat.

FT	CHAIN	18	695	FOLLICLE STIMULATING HORMONE RECEPTOR
FT	DOMAIN	18	366	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	367	387	1 (POTENTIAL).
FT	DOMAIN	388	398	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	399	421	2 (POTENTIAL).
FT	DOMAIN	422	443	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	444	465	3 (POTENTIAL).
FT	DOMAIN	466	485	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	486	508	4 (POTENTIAL).
FT	DOMAIN	509	528	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	529	550	5 (POTENTIAL).
FT	DOMAIN	551	573	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	574	597	6 (POTENTIAL).
FT	DOMAIN	598	608	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	609	630	7 (POTENTIAL).
FT	DOMAIN	631	695	CYTOPLASMIC (POTENTIAL).
FT	REPEAT	44	68	LRR 1.
FT	REPEAT	69	93	LRR 2.
FT	REPEAT	119	143	LRR 3.
FT	REPEAT	170	192	LRR 4.
FT	REPEAT	193	216	LRR 5.
FT	REPEAT	218	240	LRR 6.
FT	DISULFID	442	517	BY SIMILARITY.
FT	CARBOHYD	191	191	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	199	199	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	293	293	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	318	318	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE	695 AA;	78343 MW;	0D60A233729B5250 CRC64;

Query Match	4.7%;	Score 111.5;	DB 1;	Length 695;
Best Local Similarity	17.0%;	Pred. No. 0.56;		
Matches 84;	Conservative 75;	Mismatches 151;	Indels 183;	Gaps 16;

[illegible]

```

QY 242 NQOSDFYKIPRIEIVNKLPIVAITLLSLVY----- 271
DB 247 RSTYNLKKLP-----SLEKLVALMEASLTYPHSCCAFANRRQISEIHPICNKSILRQEV 301
QY 272 -----LAGLLAAAYQLYGTYKRRF-----PFWLETWLQCRKQLG 306
DB 302 DYMOTRGQRSSLAEDNESSYSRGFDMTYAEFDYDLQNEVDVTCSPKPDAPNPECEDILG 361
QY 307 -----LLSFFFAVYVA-----YSLCLPMRRSERLYFLNMAVYQOVHANIENS 348
DB 362 YNLRVLWIFISILAITGNIIVLVTLTTSQYKLTVP-----RFLMGNLAFAD----- 408
QY 349 WNEEVRRIEMYISFGIMSLGLSLAVTSIPSVSNALNRE-----FSFIQSTLG 399
DB 409 -----LCIGIYLLLIASVDIHYSQYHNYAIDWQTGAGCDAGFFTVFASELS 456
QY 400 ---YVALLISTFH-----VLIYGMKRAFEERYRFTPPNFV-LAL 436
DB 457 VYTLTJAITLERWHTITHAMQLDCKVHRHMASVWMGMIFAFAAALPPIFGISSYMKVSI 516
QY 437 VLPSIVILLDLQL 449
DB 517 CLPMDIDSPLSQL 529

```

RESULT 5

```

ID Y538_PASMU STANDARD; PRT; 206 AA.
AC Q9CN97;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Hypothetical protein PM0538.
GN PM0538.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: Belongs to the UPF0191 family.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

CC EMBL; AB006089; AK02622.1; -.
DR HAMAP; MF_01207; -.
DR InterPro; IPR007916; UPF0191.
DR Pfam; PF05252; UPF0191; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 7 POTENTIAL.
FT TRANSMEM 41 63 POTENTIAL.
FT TRANSMEM 76 98 POTENTIAL.
FT TRANSMEM 113 135 POTENTIAL.
FT TRANSMEM 142 164 POTENTIAL.
FT TRANSMEM 169 191 POTENTIAL.
SQ SEQUENCE 206 AA; 23653 MW; 2F277729D7F11194 CRC64;

```

Query Match 4.7%; Score 109.5; DB 1; Length 206;
 Best Local Similarity 23.8%; Pred. No. 0.18;
 Matches 56; Conservative 44; Mismatches 72; Indels 63; Gaps 13;

```

QY 228 LYSFVRDVH-----PYA-----RNQOSDFYKIPRIEIVNKLPIVAITLLSLVYLAGL 275
DB 1 MLSLFRITIHVCCGPAVLAWLWSGDSQLGADPIKEIQHFLGFSALTILLIMFILG- 59
QY 276 LAAAYQLYGTYKRRFPFWLETWLQCRKQLGLSFFFAVYVAVYSLCLPMRRSERLYFLN 335
DB 60 -----KVFFYLK-----QPQLQV---LRRALGLMAWFYVLHV-YA-----YLALE 96
QY 336 MAYQOVHANIENSWNNEEVRRIEMYISFGIMSLGLSLAVTSIPSVSNALNREFFSIQ 395
DB 97 LGYD-----FSLFVQELVNR--GYLLIGALAFLLITLMALSSWSYLR--LKMGMWFYL 146
QY 396 STLGYVALLISTFHVLIIYGMKRAFEERYRFTPPNFVALVLPSTIVILLDLQLC 450
DB 147 HOLGYVALLLGAH---YVW-----SVKNVTFSSML--YLILSIMILC 184

```

RESULT 6

```

ID FSHR_CHICK STANDARD; PRT; 693 AA.
AC P79763; Q90719;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Follicle stimulating hormone receptor precursor (FSH-R) (Folliotropin
DE receptor).
GN FSHR.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=97473503; PubMed=9332357;
RA Wakabayashi N., Suzuki A., Hoshino H., Nishimori K., Mizuno S.;
RT "The cDNA cloning and transient expression of a chicken gene encoding
RT a follicle-stimulating hormone receptor.";
RL Gene 197:121-127(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=White Leghorn;
RX MEDLINE=97057887; PubMed=8902217;
RA You S., Bridgman J.T., Foster D.N., Johnson A.L.;
RT "Characterization of the chicken follicle-stimulating hormone
RT receptor (cFSH-R) complementary deoxyribonucleic acid, and expression
RT of cFSH-R messenger ribonucleic acid in the ovary.";
RL Biol. Reprod. 55:1055-1062(1996).

```

```

CC -1- FUNCTION: Receptor for follicle stimulating hormone. The activity
CC of this receptor is mediated by G proteins which activate
CC adenylylate cyclase.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC FSH/LSH/TSH subfamily.
CC -1- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

CC EMBL; D87871; BA013487.1; -.
DR EMBL; U51097; AAC60030.1; -.
DR HSSP; P23945; 1XUN.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF00001; 7tm_1; 1.

```



```

DR Pfam; PF00560; LRR; 2.
DR PRINTS; PR00237; GPCRHRDOPSN.
DR SMART; SM00013; LRRT; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEPT_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEPT_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Phosphorylation; Repeat; Leucine-rich repeat.
FT SIGNAL 1 17
FT CHAIN 18 693
FT DOMAIN 18 366
FT TRANSMEM 367 387
FT DOMAIN 388 398
FT TRANSMEM 399 421
FT DOMAIN 422 443
FT TRANSMEM 444 465
FT DOMAIN 466 485
FT TRANSMEM 486 508
FT TRANSMEM 509 528
FT TRANSMEM 529 550
FT DOMAIN 551 573
FT TRANSMEM 574 597
FT TRANSMEM 598 608
FT TRANSMEM 609 630
FT DOMAIN 631 693
FT REPEAT 45 68
FT REPEAT 69 93
FT REPEAT 95 118
FT REPEAT 119 143
FT REPEAT 168 192
FT REPEAT 193 216
FT REPEAT 218 240
FT DISULFID 442 517
FT CARBOHYD 47 47
FT CARBOHYD 191 191
FT CARBOHYD 199 199
FT CARBOHYD 268 268
FT CONFLICT 4 4
FT CONFLICT 88 88
FT CONFLICT 140 140
FT CONFLICT 174 174
FT CONFLICT 191 191
FT CONFLICT 329 329
SQ SEQUENCE 693 AA; 78697 MW; 46F9869635A1BEC CRC64;

Query Match 4.6%; Score 109; DB 1; Length 693;
Best Local Similarity 17.2%; Pred. No. 0.87;
Matches 100; Conservative 72; Mismatches 173; Indels 238; Gaps 21;

QY 18 LPNG-INGIKARKYTVG-----VIGSGDFAKSLTIRLCRGHVIVGSRNPKFASFFP 71
Db 61 IPKGFTGLHDLKEIKISQNDALIEIGNVSSSL-----PKL----- 97

QY 72 HVDVTHEDALTKNIIFVAIHREHYTSLMDIRHLVG-----KILI 114
Db 98 -----HEIRIEKANKL-WKIDQDAFOHLPRLRYLLISNTGLSFLPVVHKVHSFOKVL 149

QY 115 DVSNNMRINQYVESNAEYLAFLPDSLIVKGFNVVSAMALQLPKASRQVYICSNMQA 174
Db 150 DVQDNHIRTIERNTFMGLSS--ESVILR-----LNKNGIOE 184

QY 175 RQGVIELARQLNFIPIDIGSLSSAREIENLPRLFTLMRGPVVAISLATFFFLYSFVRD 234
Db 185 IKD-----HAFNGTCLDEINLSDNYNLEKLPEKVFQGAIGPVVLDISRTISFLPSHGLE 239

QY 235 VIHYPARNQSDFYKPIPIETVNTKLPVIVAITLLSLVYLAGL----- 275
Db 240 FIKKLARSTYKTKLP--DYNKFRSLIEANFTYPSHCCAFTRNKTQNTTEFYPICSMSPA 297

QY 276 -----LAAAYQLYVGTKY-----RRF-----PMWLETWLQCR 302
Db 298 KQDLGEQTGRKRRRSAAEDYISHYGTRFGPVENEFDYGLCNEVVDVFCSPKDAFNPCE 357

QY 303 KQLG-----LISFFFAMVHVA-----YSICLPMRRSERYLLFNMAVQOVHAN 344

```

```

Db 358 DIMGYNVLRLVIMFINILAITGNTVLLILLISSQYKLTVP-----FILMCNLAFFAD---- 408
QY 345 IENSWNEEWRIEMYSISGIMSLGLSLAVTSIPSVSNALNRE-----FSFIQ 395
Db 409 -----LCIGIYLLFIASVDIQTKSRYYNAYIDWQGAGCNAGFTTVA 452
QY 396 STLGYVALLISTFH-----VLIYWKRAF----- 419
Db 453 SELSVYTLTVITLERMHTTYAMQLNRKYLRLHAVIIMFGMFAFTVALLPIFGISSYM 512
QY 420 -----EEERYRFTYTPNFVLAIVPSIVILDLQLCRY 452
Db 513 KVSICLPMHIETPFSQAYV--IFLLVNLVLAFFVITICICYICY 553

RESULT 7
FREQ_YEAST STANDARD; PRT; 712 AA.
ID FREQ_YEAST
AC Q12473;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ferric reductase transmembrane component 6 precursor (EC 1.16.1.7)
DE (Ferric-chelate reductase 6).
GN FREQ OR YLL051C OR L0593.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=97313267; PubMed=9169871;
RA Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansoerge W.,
RA Benes V., Brueckner M., Delius H., Dubois E., Duesterhoeft A.,
RA Eutlian K.-D., Floeth M., Goffeau A., Hebling U., Heumann K.,
RA Heuss-Neitzel D., Hilbert H., Hilger F., Miosga T., Moestl D.,
RA Louis E.J., Messenguy F., Mewes H.-W., Mosga T., Moestl D.,
RA Mueller-Auer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,
RA Portetelle D., Purnelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
RA Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz S.,
RA Underwood A.P., Urrestarazu L.A., Vandenberg M., Verhasselt P.,
RA Viereendeels F., Voet M., Volckaert G., Voss H., Wambuit R., Wedler E.,
RA Medler H., Zimmermann F.K., Zollner A., Hani J., Hohnsels J.D.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RL Nature 387:87-90(1997).
CC -1- CATALYTIC ACTIVITY: NADH + 2 Fe(3+) = NAD(+) + 2 Fe(2+).
CC -1- COFACTOR: FAD (Probable).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: Belongs to the FRE / CYBB family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z47973; CAA88006.1; -
DR EMBL; Z73156; CAA97503.1; -
DR PIR; S50969; S50969.
DR Germonline; 142046; -
DR SGD; S0003974; FREQ.
DR InterPro; IPR002916; Ferric_reduct.
DR Pfam; PF01794; Ferric_reduct_1.
KW Oxidoreductase; Electron transport; Transmembrane; Iron transport;
KW FAD; NAD; Glycoprotein; Signal; Multigene family.
FT SIGNAL 1 17
FT CHAIN 18 712
FT NP_BIND 493 499 FAD (POTENTIAL).

```

```

FT TRANSMEM 168 188 POTENTIAL.
FT TRANSMEM 245 265 POTENTIAL.
FT TRANSMEM 288 308 POTENTIAL.
FT TRANSMEM 329 349 POTENTIAL.
FT TRANSMEM 361 381 POTENTIAL.
FT TRANSMEM 388 408 POTENTIAL.
FT TRANSMEM 417 437 POTENTIAL.
FT TRANSMEM 478 498 POTENTIAL.
FT TRANSMEM 550 570 POTENTIAL.
FT TRANSMEM 553 566 POLY-LEU.
FT CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 450 450 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 655 655 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 712 AA; 81989 MW; 5224F12B51544BAA CRC64;

```

Query Match 4.6%; Score 109; DB 1; Length 712;
 Best Local Similarity 18.8%; Pred. No. 0.9;
 Matches 72; Conservative 51; Mismatches 119; Indels 142; Gaps 17;

```

QY 88 IIFVAIHREHYTSL-----WDLRHLLVGLKILIDVSNMNRINQYPSNAEYLA 136
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 183 IIAVFHMSHYNGNLRALFASRFVNIIRGHFVLPFTLVLD---KANHFKFLNVEVFTGL 238
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 137 FPDLSLVKGFNVSAVALQLGPKDASRQVYICSNNIQARQVIELARQLNFIPIDLGSL 196
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 239 MPNSL-----EAW----- 246
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 197 SARIEIENPLRLFTWRGPVVAISLATFFPL-YSFVRDVIHPYARNOQSDFYKIPPIV 255
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 247 -----IIFGYTLANTIFLSISYIIDPYNLIFNSHLSQFTRL---LA 284
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 256 NKTLPVAVITLTLVYLGLLAAAYQLYGYTKYRFPFPLWLTWLCRKQLGLSFFPAMV 315
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 285 DRS-GILAFQFPLIITFARNSPLEFLTGVKNSF-----ISFKMIGRIMVLNATI 336
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 316 H-VAYSLCLPMRSEERYFLNMAYYQVHANINENSWNEEVRIMYISGIMSLGLSL 374
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 337 HSLSYSL-----FAINHAFAK-----ISNK-----QLYKFGIASITVLCVL 373
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 375 AVTSIPSVSNALNWRSEFISQTLGYVALLI---STFVLIY-GWKRAE-----EE 422
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 374 LVLSLGIVRK---RHYEFPLYTHIILALFFECWQHVKIENGWKEMVIVSLIWGLEK 429
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 423 YRFFVTPPNFVLALVLPSTIVLDL 446
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 430 LFRIMN---ILQFRFPKATLINL 449
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 8
 FRET_YEAST STANDARD; PRT; 629 AA.

```

AC Q12333;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ferric reductase transmembrane component 7 (EC 1.16.1.7) (Ferric-
  chelate reductase 7).
GN FRET OR YOL152W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
  Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OC NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1679;
RX MEDLINE=96132030; PubMed=8553699;
RA Casamayor A., Aldea M., Casas C., Herrero E., Gamo F.J.,
  Lafuente M.J., Gancedo C., Arino J.;
RT "DNA sequence analysis of a 13 kbp fragment of the left arm of yeast
  chromosome XV containing seven new open reading frames.";
RL Yeast 11:1281-1288(1995).

```

```

CC -1- CATALYTIC ACTIVITY: NADH + 2 Fe(3+) = NAD(+) + 2 Fe(2+).
CC -1- COFACTOR: FAD (Probable).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: Belongs to the FRET / CYBB family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
-----

```

```

DR EMBL; Z48239; CAA88276.1; -.
DR EMBL; Z74894; CAA99174.1; -.
DR PIR; S60385; S60385.
DR GerMOnline; 143574; -.
DR SGD; S0005512; FRET.
DR InterPro; IPR002916; Ferric_reduct.
DR Pfam; PF01794; Ferric_reduct.1.
KW Oxidoreductase; Electron transport; Transmembrane; Iron transport;
  FAD; NAD; Glycoprotein; Multigene family.
KW NP_BIND 369 375 FAD (POTENTIAL).
FT TRANSMEM 46 66 POTENTIAL.
FT TRANSMEM 108 128 POTENTIAL.
FT TRANSMEM 168 188 POTENTIAL.
FT TRANSMEM 195 215 POTENTIAL.
FT TRANSMEM 238 258 POTENTIAL.
FT TRANSMEM 266 286 POTENTIAL.
FT TRANSMEM 293 313 POTENTIAL.
FT TRANSMEM 422 442 POTENTIAL.
FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 541 541 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 629 AA; 71996 MW; 2384480E9289C16F CRC64;

```

Query Match 4.6%; Score 107.5; DB 1; Length 629;
 Best Local Similarity 21.7%; Pred. No. 1;
 Matches 57; Conservative 46; Mismatches 77; Indels 83; Gaps 17;

```

QY 214 GPVVAISLATFFFLVSPVRDVIHPYARNOQSDFYKIPPIVNTKLPVAVITLTLVLY-L 272
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 123 GTFVVMATLTLYTLVLCFVP---HPFYR-PCAGFGSPPLSV--RAGIMVLSLVFVFSL 175
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 273 AG-----LLAAAYQLYGYTKYRFPFPLWLTWLCRKQLGLSFFPAMVAVYSLCLPM 325
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 176 SGKINIVGLVLGLSYE-----KINIVHQM-----ASILCLFFSWVHV-----IPF 215
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 326 RRSERYFLNMAYYQVHANINENSWNEEVRIMYISGIMSLGLSLAVTSIP----- 380
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 216 LRQARH---EGYERHM-----QRWKASDMWR-----SGVPILFLNLMWLSLPPIARRH 262
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 381 --SVSNALNWRSEFISQTLGYVALLISTFHV-----LIYWKRAFEERYR- 425
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 263 FYETFLQLHM-----ILAVGFYISLF---YHVPPLNSHMYLVATIVVW---FAQLFYRL 311
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 426 ----FYTPPNFVLALVLPSTIVL 444
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 312 AVKGYLRPGRSFMASTIANVSIV 334
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 9
 FSHR_HUMAN STANDARD; PRT; 695 AA.

```

AC P23945; Q16225;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Follicle stimulating hormone receptor precursor (FSH-R) (Follicotropin
  receptor).
GN FSHR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```


OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND VARIANTS ALA-307 AND SER-680.
 RC TISSUE=Ovary;
 RX MEDLINE=91222171; PubMed=1709010;
 RA Minegishi T., Nakamura K., Takakura Y., Ibuki Y., Igarashi M.;
 RT "Cloning and sequencing of human FSH receptor cDNA.";
 RL Biochem. Biophys. Res. Commun. 175:1125-1130(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=93246012; PubMed=1301382;
 RA Kelton C.A., Cheng S.V., Nugent N.P., Schweickhardt R.L.,
 RA Rosenthal J.L., Overton S.A., Wands G.D., Kuzeja J.B., Luchette C.A.,
 RA Chappel S.C.;
 RT "The cloning of the human follicle stimulating hormone receptor and
 RT its expression in COS-7, CHO, and Y-1 cells.";
 RL Mol. Cell. Endocrinol. 89:141-151(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Tilly L.T., Aihara T., Nishimori K., Jai X.-C., Billig H.,
 RA Kowalski K.I., Perlas E.A., Hsueh A.J.;
 RL Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 1-342 FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=93075197; PubMed=1359889;
 RA Gromoll J., Gudermann T., Nieschlag E.;
 RT "Molecular cloning of a truncated isoform of the human follicle
 RT stimulating hormone receptor.";
 RL Biochem. Biophys. Res. Commun. 188:1077-1083(1992).
 RN [5]
 RP SEQUENCE OF 1-51 FROM N.A.
 RX MEDLINE=95011044; PubMed=7926278;
 RA Gromoll J., Dankbar B., Gudermann T.;
 RT "Characterization of the 5' flanking region of the human follicle-
 RT stimulating hormone receptor gene.";
 RL Mol. Cell. Endocrinol. 102:93-102(1994).
 RN [6]
 RP SEQUENCE OF 286-695 FROM N.A., AND VARIANT SER-680.
 RX MEDLINE=95000244; PubMed=7916967;
 RA Gromoll J., Ried T., Holtgreve-Grez H., Nieschlag E., Gudermann T.;
 RT "Localization of the human FSH receptor to chromosome 2p21 using a
 RT genomic probe comprising exon 10.";
 RL J. Mol. Endocrinol. 12:265-271(1994).
 RN [7]
 RP 3D-STRUCTURE MODELING OF 49-228.
 RX MEDLINE=96363672; PubMed=8747461;
 RA Jiang X., Dreano M., Buckler D.R., Cheng S., Ythier A., Wu H.,
 RA Hendrickson W.A., el Tayar N.;
 RT "Structural predictions for the ligand-binding region of glycoprotein
 RT hormone receptors and the nature of hormone-receptor interactions.";
 RL Structure 3:1341-1353(1995).
 RN [8]
 RP VARIANTS ALA-307; ARG-524 AND SER-680.
 RX MEDLINE=99318093; PubMed=10391209;
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipschutz R., Daley G.Q.,
 RA Lander E.S.;
 RT "Characterization of single-nucleotide polymorphisms in coding regions
 RT of human genes.";
 RL Nat. Genet. 22:231-238(1999).
 RN [9]
 RP ERRATUM.
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipschutz R., Daley G.Q.,
 RA Lander E.S.;
 RL Nat. Genet. 23:373-373(1999).
 RN [10]
 RP VARIANTS ALA-307 AND SER-680.
 RX MEDLINE=22054685; PubMed=12059813;

RA Asatiani K., Gromoll J., Eckardstein S.V., Zitzmann M., Nieschlag E.,
 RA Simoni M.;
 RT "Distribution and function of FSH receptor genetic variants in normal
 RT men.";
 RL Andrologia 34:172-176(2002).
 RN [11]
 RP VARIANT OHSS ILE-449.
 RX MEDLINE=22812036; PubMed=12930927;
 RA Vasseur C., Rodien P., Beau I., Desroches A., Gerard C.,
 RA de Poncheville J., Chaplot S., Savagner F., Croue A., Mathieu E.,
 RA Lahlou N., Descamps P., Misrahi M.;
 RT "A chorionic gonadotropin-sensitive mutation in the
 RT follicle-stimulating hormone receptor as a cause of familial
 RT gestational spontaneous ovarian hyperstimulation syndrome.";
 RL New Engl. J. Med. 349:753-759(2003).
 RN [12]
 RP VARIANT OHSS ASN-567.
 RX MEDLINE=22812037; PubMed=12930928;
 RA Smits G., Olatundusun O., Delbaere A., Pierson R., Vassart G.,
 RA Costagliola S.;
 RT "Ovarian hyperstimulation syndrome due to a mutation in the
 RT follicle-stimulating hormone receptor.";
 RL New Engl. J. Med. 349:760-766(2003).
 CC -!- FUNCTION: Receptor for follicle stimulating hormone. The activity
 CC of this receptor is mediated by G proteins which activate
 CC adenylylate cyclase.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Name=Long;
 CC IsoId=P23945-1; Sequence=Displayed;
 CC Name=Short;
 CC IsoId=P23945-2; Sequence=VSP_001953, VSP_001954;
 CC -!- TISSUE SPECIFICITY: Sertoli cells and ovarian granulosa cells.
 CC -!- DISEASE: Defects in FSHR are a cause of ovarian hyperstimulation
 CC syndrome (OHSS) [MIM:608115]. OHSS is a disorder which occurs
 CC either spontaneously or most often as an iatrogenic complication
 CC of ovarian stimulation treatments for in vitro fertilization. The
 CC clinical manifestations vary from abdominal distention and
 CC discomfort to potentially life-threatening, massive ovarian
 CC enlargement and capillary leak with fluid sequestration.
 CC Pathologic features of this syndrome include the presence of
 CC multiple serous and hemorrhagic follicular cysts lined by
 CC luteinized cells, a condition called hyperreactio luteinalis.
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC FSH/LSH/TSH subfamily.
 CC -!- SIMILARITY: Contains 5 leucine-rich (LRR) repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M65085; AAA52477.1; -;
 DR EMBL; S59900; AAB26480.1; -;
 DR EMBL; M95489; AAA52478.1; -;
 DR EMBL; X68044; CAA48179.1; -;
 DR EMBL; S73199; AAB32071.1; -;
 DR EMBL; S73526; AAB32225.1; -;
 DR PIR; I57661; QRHUF1.
 DR PDB; 1XUN; 15-MAY-97.
 DR Genew; HGNC:3969; FSHR.
 DR MIM; 136435; -;
 DR MIM; 608115; -;
 DR GO; GO:0016021; C:integral to membrane; TAS.
 DR GO; GO:0005886; C:plasma membrane; TAS.
 DR GO; GO:0004963; F:follicle stimulating hormone receptor activity; TAS.
 DR GO; GO:0007292; P:female gamete generation; TAS.
 DR GO; GO:0008585; P:female gonad development; TAS.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signaln. . .; TAS.

```

DR GO; GO:0007283; P:spermatogenesis; TAS.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR; 3.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR SMART; SM00013; LRRNT; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing;
KW 3D-structure; Polymorphism; Disease mutation.
FT SIGNAL 1 17
FT CHAIN 18 695
FT DOMAIN 18 366 FOLLICLE STIMULATING HORMONE RECEPTOR.
FT TRANSMEM 367 387 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 388 398 1 (POTENTIAL).
FT TRANSMEM 399 421 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 422 443 2 (POTENTIAL).
FT TRANSMEM 444 465 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 466 485 3 (POTENTIAL).
FT TRANSMEM 486 508 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 509 528 4 (POTENTIAL).
FT TRANSMEM 529 550 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 551 573 5 (POTENTIAL).
FT TRANSMEM 574 597 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 598 608 6 (POTENTIAL).
FT TRANSMEM 609 630 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 631 695 7 (POTENTIAL).
FT REPEAT 69 93 CYTOPLASMIC (POTENTIAL).
FT REPEAT 119 143 LRR 1.
FT REPEAT 170 192 LRR 2.
FT REPEAT 193 216 LRR 3.
FT REPEAT 218 240 LRR 4.
FT DISULFID 442 517 LRR 5.
FT CARBOHYD 191 191 BY SIMILARITY.
FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 318 318 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 224 285 Missing (in isoform Short).
FT /FTid=VSP_001953.

```

Query Match 4.6%; Score 107.5; DB 1; Length 695;
 Best Local Similarity 17.0%; Pred. No. 1.1;
 Matches 84; Conservative 77; Mismatches 148; Indels 185; Gaps 17;

```

QY 79 HEDALTKNIIIFVAIHREHYTSLMDLRHLVG-----KILDIVSNMR 121
DB 98 HEIRIEKANNL-LYINPEAFQNLNLYLLISNTGIKHLPDVHKHSLQKYLDDIQDNIN 156
QY 122 INQYVESNAEYLAFLPDSLLVKGFNVSAMALQLGPKDASRQVYICSNNTQARQVIEL 181
DB 157 IH-----TIERNSFVGLSFESVILWL-----NKGIOEIHNCA----- 189
QY 182 ARQLNFIPIIDGLSSAREIENLPLRLFTLWRGPNVVAISLATFFLYSPFVDVIHPYAR 241
DB 190 ---FNGTQLDELNLSDNNNLIELPNDVFHGASGPVILDISRTRIHSLPSGLENLKLLRA 246
QY 242 NQOSDFYKIPRIBIVNKL-PIVAITLLSLV-----PPULETWLQCRKQL 305
DB 247 RSTYNLKLKLP-----TLELVALMEASLTYPHSCCAFANWRQISELHPICNKSILROE 300
QY 272 -----LAGLLAAAYQLYYGTYKRRF-----PPULETWLQCRKQL 305
DB 301 VDYMTQTRGQRSSLAEDNESYSRGFDMTYTEFDYDLCEVVDVTCSPKPDAPNCEDIM 360
QY 306 G-----LISFFPAMVHA-----YSLCLPMKRSEERYLFLNMAVQOVHANIEN 347
DB 361 GYNILRLVLIWFISITAITGNILVILVLTTSQYKLTVP-----RFLMCNLAFAD----- 408
QY 348 SWNEEEVWRIEMYISFGIMSLGLSLLLAVTSIPSVSNALNMR-----FSFIQSTL 398

```

```

DB 409 -----LCIGIYLLLIASVDIHTKSQYHNVAIDWQTAGCDAGFTVFASEL 455
QY 399 G---YVALLISTFH-----VLIYWKRAFEERYRFTPPNFV-LA 435
DB 456 SVYTLFAITLERHTITHAMQLDCKVQLRHAASVWVMGMIFAFAALFPFIIGISSYMKVS 515
QY 436 LVLPSIVILLDLQL 449
DB 516 ICLFMDIDSPLSQL 529

```

```

RESULT 10
FSHR_PIG STANDARD; PRT; 695 AA.
ID FSHR_PIG
AC P49059; 077514;
DT 01-FEB-1996 (Rel. 33, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Follicle stimulating hormone receptor precursor (FSH-R) (Follictrophin
DE receptor).
GN FSHR.
OS Sus scrofa (pig).
OC Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eukaryota; Metazoa;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_Taxid=9623;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=96011644; PubMed=7590277;
RA Remy J.J., Lahbib-Mansals Y., Yerle M., Bozon V., Couture L.,
RA Pajot E., Grebert D., Saesle R.;
RT "The porcine follicotropin receptor: cDNA cloning, functional
RT expression and chromosomal localization of the gene.";
RL Gene 163:257-261(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Wang Y.F., Meyer K.B., Schmidt K., Wan S.J., Degen S.J.F.,
RA la Barbera A.R.;
RT "Porcine follicle-stimulating hormone receptor.";
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Receptor for follicle stimulating hormone. The activity
CC of this receptor is mediated by G proteins which activate
CC adenylylate cyclase.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC FSH/LSH/TSH subfamily.
CC -!- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L31966; AAA86933.1; -.
DR EMBL; AF025377; AAC24981.1; -.
DR HSSP; P23945; 1XUN.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR; 4.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR SMART; SM00013; LRRNT; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Phosphorylation; Repeat; Leucine-rich repeat.

```



```

FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 695 FOLLICLE STIMULATING HORMONE RECEPTOR.
FT DOMAIN 18 366 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 367 387 1 (POTENTIAL).
FT DOMAIN 388 398 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 399 421 2 (POTENTIAL).
FT DOMAIN 422 443 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 444 465 3 (POTENTIAL).
FT DOMAIN 466 485 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 486 508 4 (POTENTIAL).
FT DOMAIN 509 528 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 529 550 5 (POTENTIAL).
FT DOMAIN 551 573 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 574 597 6 (POTENTIAL).
FT DOMAIN 598 608 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 609 630 7 (POTENTIAL).
FT DOMAIN 631 695 CYTOPLASMIC (POTENTIAL).
FT REPEAT 44 68 LRR 1.
FT REPEAT 69 93 LRR 2.
FT REPEAT 119 143 LRR 3.
FT REPEAT 170 192 LRR 4.
FT REPEAT 193 216 LRR 5.
FT REPEAT 218 240 LRR 6.
FT DISULFID 442 517 BY SIMILARITY.
FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 2 2 S -> A (IN REF. 1).
FT CONFLICT 13 13 T -> S (IN REF. 1).
FT CONFLICT 60 60 V -> A (IN REF. 1).
FT CONFLICT 166 166 V -> M (IN REF. 1).
FT CONFLICT 215 215 Q -> H (IN REF. 1).
FT CONFLICT 247 247 K -> R (IN REF. 1).
FT CONFLICT 257 257 S -> T (IN REF. 1).
FT CONFLICT 334 334 D -> N (IN REF. 1).
FT CONFLICT 349 349 E -> K (IN REF. 1).
FT CONFLICT 352 352 T -> A (IN REF. 1).
FT CONFLICT 383 383 V -> E (IN REF. 1).
FT CONFLICT 407 407 A -> T (IN REF. 1).
FT CONFLICT 421 421 V -> I (IN REF. 1).
FT CONFLICT 427 427 T -> S (IN REF. 1).
FT CONFLICT 435 435 D -> N (IN REF. 1).
FT CONFLICT 483 483 L -> V (IN REF. 1).
FT CONFLICT 550 550 T -> I (IN REF. 1).
FT CONFLICT 586 586 A -> V (IN REF. 1).
FT CONFLICT 607 607 S -> L (IN REF. 1).
FT CONFLICT 691 691 R -> H (IN REF. 1).
SQ SEQUENCE 695 AA; 78172 MM; E9EBEDB29C79C450 CRC64;

```

```

Query Match 4.4%; Score 104.5; DB 1; Length 695;
Best Local Similarity 16.5%; Pred. No. 1.9;
Matches 91; Conservative 91; Mismatches 169; Indels 199; Gaps 21;

```

```

Qy 18 LPNG-INGIKDARKVTGVTGSGDPAKSLTIRLCRGYHVIGSRNPKFASFPFHVVDV 76
Db 61 IPKGAFSGFDLEKI-----EISQNDVLEVEAN---VFSNLPKL----- 97
Qy 77 THHEDALTKNIIFVAIHREHYTSLMDLRHLVG-----KILIDVSNM 119
Db 98 --HEIRIEKANNL-LYIDPDAPQNLPRKRYLLISNTGVKHLPAVHKIQSLQKVLDDIQDN 154
Qy 120 MRINQYPSNAEYLAFLPDSLIVGKFNVSAMALQGPKDSRQVYICSNNTQARQOVI 179
Db 155 INIH-----TVERNSFVGSGFESMILWL---SKNGIREIHNC----- 189
Qy 180 ELARQLNFIPIDLGSLSSAREIENLPRLFTLWRGPVVAISLATFFLYSFVRDVTHPY 239
Db 190 ----FNGTQDELNLSNDNLEELPDVDFQASGPVILDISRTRIHSLPSYGLENIKKL 244
Qy 240 ARNQSDPYKIP-----IEIVNKT-----PIVAITL----- 267
Db 245 RAKSTYNLKLPSELEKFTVLTMEASLTYPESHCCAFANMRRQISDLHPICNKSILRQEVDM 304

```

```

Qy 268 ----SLVYLAGLLAAVQLYGTYRRF-----PWLFTWLCRKQLG--- 306
Db 305 TQARGQVSLAEDGESSLAKEPDTMYSEFDYDLCEVVDVICSPEPDTFNPCEIDMGHDI 364
Qy 307 --LISFFPAMVHA-----YSLCLPMRBSERYLFLNMAYQOVHANIENSWNE 351
Db 365 LRVLIWFISILAITGNIVLVILITSQYKLTVP-----RFLMCLAFAD----- 408
Qy 352 EEVWRIEMYISFGIMSLGLSLLAVTSIPSVSNALNWR-----FSFIQSTLG--- 399
Db 409 -----LCIGIYLLIASVDIHTKTQYHNVAIDWQTGAGCDAGFTVFASELVYT 459
Qy 400 YVALLISTFH-----VLIGWKRAFEERYRFTPPNFV-LALVLP 439
Db 460 LTAITLERWHTITHAMQLOCKVQLRHAASIMLVGWIPTVALPFIPIGSISSYMKVSIQLP 519
Qy 440 SIVLDLLOL 449
Db 520 MDIDSPLSQL 529

```

```

RESULT 11
ID SOTB_EC057 STANDARD; PRT; 396 AA.
AC P58529;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Sugar efflux transporter.
GN SOTB OR Z2173 OR ECS2135.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RL Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22(2001).
CC -!- FUNCTION: Involved in the efflux of sugars. The physiological role
CC may be the reduction of the intracellular concentration of toxic
CC sugars or sugar metabolites. Transports L-arabinose and to a
CC lesser extent IPTG. Seems to contribute to the control of the
CC arabinose regulon (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Probable).
CC -!- SIMILARITY: Belongs to major facilitator superfamily. SotB
CC (TC 2.A.1.2) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE005352; AAG56236.1; -.
DR EMBL; AP002557; BAB35558.1; -.
DR PIR; G90895; G90895.
DR PIR; H85721; H85721.
DR HAMAP; MF_00517; -.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR PROSITE; PS50850; MFS; 1.
KW Transport; Sugar transport; Transmembrane; Inner membrane;
KW Complete proteome.
FT TRANSMEM 15 35 POTENTIAL.
FT TRANSMEM 50 70 POTENTIAL.
FT TRANSMEM 81 101 POTENTIAL.
FT TRANSMEM 103 123 POTENTIAL.
FT TRANSMEM 136 156 POTENTIAL.
FT TRANSMEM 170 190 POTENTIAL.
FT TRANSMEM 209 229 POTENTIAL.
FT TRANSMEM 246 266 POTENTIAL.
FT TRANSMEM 275 295 POTENTIAL.
FT TRANSMEM 299 319 POTENTIAL.
FT TRANSMEM 333 353 POTENTIAL.
FT TRANSMEM 364 384 POTENTIAL.
SQ SEQUENCE 396 AA; 42568 MW; CB6A209B0BE6D4F0 CRC64;

Query Match
Best Local Similarity 4.4%; Score 104; DB 1; Length 396;
Matches 85; Conservative 72; Mismatches 148; Indels 146; Gaps 20;

QY 31 VTVGVIGSGDFAKSLTIRLCRGYHVVIGSRNPKFASFEFPHVVDVTHEDALTKNIF 90
Db 31 VPVGLL--SDIAQSFHMOTAQVGMILTIVAMVVALMSLPFMLTSGVERKLLICLFVVF 88
QY 91 VAIHREHYTSLWDLHLVGLKILIDVSNMNRINQYPESNAEYLASLPEDSLIVKGFNVVS 150
Db 89 IASHVLSFLS-WSFTVLVISRI-----GVAFAHAIF-----WSITA 123
QY 151 AMALQLGP--KDSARQVYICSNNIQARQVIELAR-----QINFPIDLSL----- 195
Db 124 SLAIRMAPAKRAQALSLIATGTALAMVGLPLGRIVGOYFGWRMTFPAIGIGALITLLC 183
QY 196 -----SSAREIENLPLRLFTLWRGPVVAISLAT-----FFFLYSFVRDVIHPY 239
Db 184 LIKLPLLPSEHSGSLKSLPL----LFRPALMSIYLLTVVVTATAYASY----LEPF 235
QY 240 ARN---QQSDFYKPIREIVNKTLPVIAITLGLVYLAGLLAAYQLYGKRYRFPFWLE 296
Db 236 VQNIAGFSANF-----ATALLLLGAGIIGSVIFGKLGNOYAS----- 274
QY 297 TWLQCRKQLGLSFPFAMVAVVAVSLCLPMRSEKYLFLMAYQGVHANTENSMNEEVR 356
Db 275 -----ALVSTAILLLVCLALLPANSF-----IHLGLSIF-----WG 309
QY 357 IEMYISFGIMSLGLSL-----LAVTSIPSVSN-----ALNWRFSFIQST 397
Db 310 IAMMITIGLMQVKVLALADATDVAMALFSGIFNIGIGAGALVGNQVSLHMS-----MSM 364
QY 398 LGYVALLISTFHVLIYG-----WKRAFEFE 422
Db 365 IGYVG-TVPAPALIMSIIFRRWPVTLLEQ 394

RESULT 12
SOTB_ECOLI STANDARD; PRT; 396 AA.
AC P31122; P76883; P77353;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Sugar efflux transporter.
GN SOTB OR B1528.

```

```

OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxId=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itch T.,
RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
RA Sempel G., Seki Y., Siyasundaram S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
RN [3]
RP SEQUENCE OF 217-396 FROM N.A.
RX MEDLINE=93186717; PubMed=8383113;
RA Cohen S.P., Haechler H., Levy S.B.;
RT "Genetic and functional analysis of the multiple antibiotic
resistance (mar) locus in Escherichia coli.";
RL J. Bacteriol. 175:1484-1492(1993).
RN [4]
RP CHARACTERIZATION.
RC STRAIN=SB0;
RX MEDLINE=99194728; PubMed=10094697;
RA Bost S., Silva F., Belin D.;
RT "transcriptional activation of ydeA, which encodes a member of the
major facilitator superfamily, interferes with arabinose accumulation
and induction of the Escherichia coli arabinose PBAD promoter.";
RL J. Bacteriol. 181:2185-2191(1999).
RN [5]
RP CHARACTERIZATION.
RC STRAIN=J5219;
RX MEDLINE=99369894; PubMed=10438792;
RA Carole S., Pichoff S., Bouche J.-P.;
RT "Escherichia coli gene ydeA encodes a major facilitator pump which
exports L-arabinose and isopropyl-beta-D-thiogalactopyranoside.";
RL J. Bacteriol. 181:5123-5125(1999).
CC -I- FUNCTION: Involved in the efflux of sugars. The physiological role
may be the reduction of the intracellular concentration of toxic
sugars or sugar metabolites. Transports L-arabinose and to a
lesser extent IPTG. Seems to contribute to the control of the
arabinose regulon.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
(Possible).
CC -I- SIMILARITY: Belongs to major facilitator superfamily. SotB
(TC 2.A.1.2) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000250; AAC74601.1; -.
DR EMBL; D90795; BAA15210.1; -.
DR EMBL; D90796; BAA15218.1; -.
DR EMBL; D90797; BAA15230.1; -.

```



```

DR EMBL; M96235; -; NOT_ANNOTATED_CDS.
DR PIR; C64907; C64907.
DR Ecogene; EG11636; scb.
DR HAMAP; MF_00517; -; 1.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR Pfam; PF00083; sugar tr; 1.
DR PROSITE; PS50850; MFS; 1.
DR Transport; Sugar transport; Transmembrane; Inner membrane;
KW Complete proteome.
FT TRANSMEM 15 35 POTENTIAL.
FT TRANSMEM 50 70 POTENTIAL.
FT TRANSMEM 81 101 POTENTIAL.
FT TRANSMEM 103 123 POTENTIAL.
FT TRANSMEM 136 156 POTENTIAL.
FT TRANSMEM 170 190 POTENTIAL.
FT TRANSMEM 209 229 POTENTIAL.
FT TRANSMEM 246 266 POTENTIAL.
FT TRANSMEM 275 295 POTENTIAL.
FT TRANSMEM 299 319 POTENTIAL.
FT TRANSMEM 333 353 POTENTIAL.
FT TRANSMEM 364 384 POTENTIAL.
SQ SEQUENCE 396 AA; 42538 MW; CB6A34CA4EB6D4F0 CRC64;

Query Match
Best Local Similarity 18.4%; Score 104; DB 1; Length 396;
Matches 85; Conservative 72; Mismatches 148; Indels 146; Gaps 20;

QY 31 VTGVIGSGDFAKSLTIRLCGYHVVIGSRNPKFASEFPFVVDVTHEDALTKNIIF 90
DB 31 VPVGLL--SDIAQSFHMOTAQVGIMLTYYAWVALMSLPFMLMTSQVERKLLICLVVF 88
QY 91 VAIHREHYSLMDLRHLVGLKILIDVSNMNRINQYPESNAEYIASLPDLSLVKGFNVVS 150
DB 89 IASHVLSFLS--WSFTVLVISRI-----GVAFAHAIF-----WSITA 123
QY 151 AWALQLGP--KDSARQVYICSNINQARQVIELAR-----QLNFIPIDLSL----- 195
DB 124 SLAIRMAPAKRAQALSLIATGTALAVLGLPLGRIVGQYFGWRMTFFAIGALITLLC 183
QY 196 -----SSAREIENLPLRLFTLRGPPVVAISLAT-----FFFLYSFVRDVVHPY 239
DB 184 LKLLPLRSEHSGSLKSLPL---LFRRPALMSIYLLTVVVVTAHYTASY---LEPF 235
QY 240 ARN---QQSDFYKIPIEIVNKTLPVIAITLLSLVYLAGLLAAVQLYYGTKYRRFPWLE 296
DB 236 VQNIAGFSANF-----ATALLLLLGAGIIGSVIFGKLGNOYAS----- 274
QY 297 TWLQCRKQGLLSFFFAMVHVAVSLCPMRSEERYLFLNMAYQOVHANTENSMNEEVR 356
DB 275 -----ALVSTAIALLLVCLALLPRANSE-----IHGLVLSIF-----WG 309
QY 357 IEMYISFGIMSLGLSL-----LAVTSIPSVSN-----ALNWRFSFIQST 397
DB 310 IAMMIIGLGMQVYVLAAPDATDVAMALRSGIFNIGIGAGALVGNQVSLHMS-----MSM 364
QY 398 LGYVALLISTFHYLYG-----WKRAFEFE 422
DB 365 IGYVG-AVPAPFAALISIIIFRRWPVTLREQ 394

```

```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN 1)
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=95335846; PubMed=7611445;
RA Chong S.S., Kozak C.A., Liu L., Kristjansson K., Dunn S.T.,
RA Bourdeau J.E., Hughes M.R.;
RT "Cloning, genetic mapping, and expression analysis of a mouse renal
RT sodium-dependent phosphate cotransporter.";
RL Am. J. Physiol. 268:F1038-F1045(1995).
CC -1- FUNCTION: Important for the resorption of phosphate by the kidney.
CC May be involved in actively transporting phosphate into cells via
CC Na(+) cotransport in the renal brush border membrane.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Kidney.
CC -1- SIMILARITY: Belongs to the sodium/phosphate cotransporter family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X77241; CAA54459.1; -.
DR PIR; S69915; S69915.
DR MGD; MGI:103209; Slc17a1.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR004745; Pi_cotranspt.
DR TIGRFAMs; TIGR00894; 2A0114euk; 1.
DR PROSITE; PS50850; MFS; 1.
DR Transport; Symport; Sodium transport; Transmembrane; Glycoprotein.
KW TRANSMEM 79 99 POTENTIAL.
FT TRANSMEM 117 137 POTENTIAL.
FT TRANSMEM 176 196 POTENTIAL.
FT TRANSMEM 199 219 POTENTIAL.
FT TRANSMEM 260 280 POTENTIAL.
FT TRANSMEM 304 324 POTENTIAL.
FT TRANSMEM 337 356 POTENTIAL.
FT TRANSMEM 363 383 POTENTIAL.
FT TRANSMEM 399 419 POTENTIAL.
FT TRANSMEM 429 449 POTENTIAL.
FT TRANSMEM 47 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 47 47 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 56 56 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 465 AA; 51589 MW; C67EE25A2C291EEF CRC64;

Query Match
Best Local Similarity 18.7%; Score 103; DB 1; Length 465;
Matches 76; Conservative 64; Mismatches 142; Indels 124; Gaps 16;

QY 110 GKILIDVSNMNRINQYPESNAEYIASLPDLSLVKGFNVSAVALQGPKDA---SRQVY 166
DB 79 GLILSSVFFGMVVQAP--VGYLSGIYPMKRIIGSSLFLSSLSMLIPPAQVGALVI 135
QY 167 ICSNNIQARQOVIELARQLNFI---PIDGLSSSAREIENLPLRLFTLRGPPVV--- 218
DB 136 VCRVLQGIAGTIVSTGQHEIWKMAPPLEGRILTS-----MTLSGFVM--GPFIVLVVS 187
QY 219 -----AISLATPFLYSFVRDVVHPYARQNSDF----- 247
DB 188 GFICDLLGWPMVRYIFGIVGCVLSLSWFFLPDDPKD--HPYMSSEKDYIISSLMQAS 245
QY 248 ---YKIPIEIVNKTLPVIAITLS----- 268
DB 246 SGRQSLPIKAKLSPLMAIILNSFAFIWNSLLVTTYPTFISTVLHVNVRENGSLSLP 305
QY 269 --LVYLAGLLAAVQLYYGTKYRRFPWLETWLQCRKQGLLSFFPAMVHVAVSLCLPMR 326

```

```

RESULT 13
NPT1 MOUSE
ID NPT1 MOUSE STANDARD; PRT; 465 AA.
AC Q61983;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Renal sodium-dependent phosphate transport protein 1 (sodium/phosphate
DE cotransporter 1) (Na(+)/Pi cotransporter 1) (Renal sodium-phosphate
DE transport protein 1) (Renal Na(+)-dependent phosphate cotransporter
DE 1).
GN SLCl7A1 OR NPT1.

```

```

Db      306 YLLAYICGILAGQMSDFELLTRK-----IFSIVTRKLEFTLLGSCFPVIFIMCLLYLSYN 359
QY      327 RSERYLFLNMAYQOVHANIEINSWNEEVRWRIEM---YISF--GIMSL-----GLLSLAV 376
Db      360 FYSTVIFLTLA---NSTLSFSYCGQLINALDAPRYVYGLKAVTALIGMGGLISSTLA 415
QY      377 TSIPSVSNALNWRPESFIQSTLGVALLISTFHYL-----IYGWR 417
Db      416 GLINQDPEYAWHKISFLMAGINVTCLV---FYFLFAKGEIQDWAK 458

RESULT 14
NUCC_NEPOL STANDARD; PRT, 391 AA.
ID NUCC_NEPOL
AC Q9TAV6;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE NAD(P)H-quinone oxidoreductase chain H, chloroplast (EC 1.6.5.-)
DE (NAD(P)H dehydrogenase, chain H) (NADH-plastoquinone oxidoreductase
DE 49 kDa subunit).
GN NDH.
OS Nephroselmis olivacea.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Chlorophyta; Prasinophyceae;
OC Chlorodendales; Chlorodendraceae; Nephroselmis.
OX NCBI_TaxID=31312;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIES-484;
RX MEDLINE=99398694; PubMed=10468594;
RA Turmel M., Otis C., Lemieux C.;
RT "The complete chloroplast DNA sequence of the green alga Nephroselmis
RT olivacea: insights into the architecture of ancestral chloroplast
RT genomes.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:10248-10253(1999).
CC -!- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
CC TO BE PLASTOQUINONE. COMPONENT OF THE IRON-SULFUR (IP) FRAGMENT OF
CC THE ENZYME.
CC -!- CATALYTIC ACTIVITY: NAD(P)H + plastoquinone = NAD(P)(+) +
CC plastoquinol.
CC -!- SIMILARITY: Belongs to the complex I 49 kDa subunit family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF137379; AAD54891.1; -
CC InterPro; IPR001135; Oxidored_49Kd.
CC Pfam; PF00346; complex1_49Kd; 1.
CC PROSITE; PS00535; COMPLEX1_49K; 1.
CC Oxidoreductase; NAD; NADP; Quinone; Plastoquinone; Chloroplast.
CC SEQUENCE 391 AA; 44913 MW; 9A40AEC68995825E CRC64;
CC -----

Query Match 4.4%; Score 102.5; DB 1; Length 391;
Best Local Similarity 17.9%; Pred. NO. 1.3;
Matches 80; Conservative 70; Mismatches 124; Indels 173; Gaps 20.

QY      85 KTNIFVAIHREHYTSLMDLRHLV--GKLLD-----VSNMRINOYPE 127
Db      5 KTDPMIVSM-GPAHPSMHGVLRLIVTLGDENVLDCEPVGYLHRGMEKTAENRTIVQYLP 63
QY      128 --SNAEYLASLFPDSLIVKGFNVSAWALQLGPKDASRQVYICSNNIQARQCVIELAROL 185
Db      64 YVTRWDYLATMTFTAITV-----NAPERLANIEVPRRA 96
QY      186 NFIPIDGLSSARETENLPLRLFTLMRGPVVAISLATEFFLYSFVRDVTHPYARNQOS 245

```

```

Db      97 SYLRVIMLELSR-----IASHLIMLGPFMADLGAQTPEFYILREREMY----- 140
QY      246 DFYKIPRIEIVNKTLPVAILTLISLVYLAGLLAAAYQLYGTKYRRFPFWLETWLQ-CRKQ 304
      141 DLFE-----AATGMKMHNYFRVGVAA--DVPG-----WIDKCLDFCEYF 180
QY      305 LGLLSFFFAMV-----HVAYSLLCPM-----RRSERYLFL 334
      181 LPKVDEYEALITRNPILFKRVKGVGTISPQOAINWGLSGPMLRAGSVMDLRKVDRY--- 237
QY      335 NMAVQOVHANIEINSWNEEVR-----EMYISFGIMSLGLSLIAVTSIP----- 380
      238 -ECYEDFHMSVESEETGDCIARYLVRIREMRSTKIVQ-----QALKSIPEGPTENLEA 290
QY      381 ---SVSNALNMRREFS-----IOSTLGYVAL-LISTFHVLIYGM 415
      291 RQLSQGRTSPWNEFDYQFLGKKASPTFKWPRQEHYVRVEAPKGELGVFLIGDDHVFPPRW 350
QY      416 KRAFEEEYRFTYTPNFVALVLPSTV 442
      351 K-----IRPPGFINVQILPENV 367

```

```

RESULT 15
Y304 BRUME
ID Y304 BRUME STANDARD; PRT; 220 AA.
AC Q8YD73;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein EMEII0304.
GN EMEII0304.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=16M / ATCC 23456 / Biotype 1;
RX MEDLINE=20020109; PubMed=11756688;
RA Delvecchio V.G., Kapatal V., Redkar R.U., Patra G., Mujer C., Los T.,
RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
RA Selkov E., Elizer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselkorn R., Kyrpides N., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis."
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
CC -I- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -I- SIMILARITY: Belongs to the UPF0191 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE009669; AAL53546.1; -.
DR PIR; AG3547; AG3547.
DR HAMAP; MF_01207; -, 1.
DR InterPro; IPR007916; UPF0191.
DR Pfam; PF05252; UPF0191; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
KM TRANSMEM 20 39 POTENTIAL.
FT TRANSMEM 54 72 POTENTIAL.
FT TRANSMEM 85 104 POTENTIAL.
FT TRANSMEM 124 146 POTENTIAL.
FT TRANSMEM 153 175 POTENTIAL.
FT TRANSMEM 179 198 POTENTIAL.
SQ SEQUENCE 220 AA; 24815 MW; 182C0244743B17FA CRC64;

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 12:37:02 ; Search time 21 Seconds
(without alignments)
2079.569 Million cell updates/sec

2079.569 Million cell updates/sec

Title: US-09-455-486-6

Perfect score: 2351

Sequence: 1 MESISMWSEKSLSETCLPN.....ALVPSIVLDLLQLCRYPD 454

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	160	6.8	198	2	D95285 conserved hypothet
2	159.5	6.8	211	2	AC2560 hypothetical prote
3	156.5	6.7	239	2	T50571 probable oxidoredu
4	147	6.3	213	2	H69400 conserved hypothet
5	142	6.0	212	2	D69361 conserved hypothet
6	142	6.0	224	2	T10120 F420-dependent NAD
7	133	5.7	223	2	D64487 hypothetical prote
8	121	5.1	191	2	F86826 hypothetical prote
9	120.5	5.1	232	2	A69131 conserved hypothet
10	119.5	5.1	216	2	T00121 conserved hypothet
11	116.5	5.0	242	2	G82642 conserved hypothet
12	116	4.9	198	2	AB3182 conserved hypothet
13	111.5	4.7	695	1	JN0898 foliitropin recept
14	111.5	4.7	1228	2	S59681 probable membr
15	110.5	4.7	222	2	B84410 hypothetical prote
16	110.5	4.7	694	2	JC4301 foliitropin recept
17	109	4.6	320	2	T28379 ORF MSV218 hypothe
18	109	4.6	712	2	S50969 probable membr
19	109	4.6	1184	2	H71436 hypothetical prote
20	109	4.6	1301	2	D85188 disease resistance
21	107.5	4.6	629	2	S60385 probable membr
22	107.5	4.6	695	1	QRHUT foliitropin recept
23	105	4.5	442	2	B64582 sodium- and chlori
24	105	4.5	604	2	T31042 hypothetical prote
25	104	4.4	396	1	C64907 chloramphenicol re
26	104	4.4	396	2	G90895 hypothetical prote
27	104	4.4	396	2	H85721 probable resistanc
28	103	4.4	465	2	S69915 sodium-phosphate t
29	102.5	4.4	314	2	AB2972 hypothetical prote

30	102.5	4.4	314	2	H98310 probable peptide A
31	102.5	4.4	1242	2	T39453 probable mrna stab
32	102	4.3	220	2	AG3547 bicyclomycin resis
33	102	4.3	574	2	T41068 hypothetical prote
34	101	4.3	320	2	E71139 hypothetical prote
35	101	4.3	348	2	T12284 NADH2 dehydrogenas
36	101	4.3	420	2	F69144 O-antigen transpor
37	101	4.3	501	2	T02134 hypothetical prote
38	101	4.3	503	2	C86250 hypothetical prote
39	101	4.3	735	2	AB3006 hypothetical prote
40	100.5	4.3	346	2	T11181 NADH2 dehydrogenas
41	100.5	4.3	1780	2	A85045 probable glucan sy
42	100	4.3	395	2	C71219 hypothetical prote
43	100	4.3	476	2	A28439 endonuclease Scel
44	100	4.3	714	2	AF2479 ABC transporter At
45	99.5	4.2	452	2	C71391 NADH2 dehydrogenas

ALIGNMENTS

RESULT 1		D95285	conserved hypothetical protein Sma0349 [imported] - Sinorhizobium meliloti (strain 1021)
		C/Species:	Sinorhizobium meliloti
		C/Date:	24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
		C/Accession:	D95285
		R/Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows	
		.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.	
		Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001	
		A/Title:	Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot
		A/Reference number:	A95262; MUID:21396509; PMID:11481432
		A/Accession:	D95285
		A/Status:	preliminary
		A/Molecule type:	DNA
		A/Residues:	1-198 <KUR>
		A/Cross-references:	GB:AB006469, PIDN:AAK64846.1, PID:914523260, GSPDB:GN00165
		A/Experimental source:	strain 1021, megaplasmid pSymA
		R/Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,	
		pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;	
		L.; Hyman, R.W.; Jones, T.	
		Science 293, 668-672, 2001	
		A/Authors:	Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
		hebaull, P.; Vandenbol, M.; Vorholter, F.U.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.	
		A/Title:	The composite genome of the legume symbiont sinorhizobium meliloti.
		A/Reference number:	A96039; MUID:21368234; PMID:11474104
		A/Contents:	annotation
		C/Genetics:	
		A/Gene:	Sma0349
		A/Genome:	plasmid
		Query Match	6.8%; Score 160; DB 2; Length 198;
		Best Local Similarity	26.7%; Pred. No. 2.1e-05;
		Matches	55; Conservative 40; Mismatches 83; Indels 26; Gaps 8;
QY	32	TVGVIGSGDFAKSLTIRLCGYHVT-GSRNPKFAS---EFFPHVVDVTHEDALTKTN 87	
DB	3	TYAIIAGAGIGSALAEFTTAQIPATIANRGPASLSSVTRFGASVAVELKDAL-QAD 61	
QY	88	IIFVAIHREHYTSLMDRLHLV--GKILIVSNMRINQY-----ESNAEYLASLFP 138	
DB	62	VVILAV--PYDSIADIVTQVSDWGGQIVVDASNAIDPPAFKPRDLGRLSTEIVSELP 118	
QY	139	DSLIVKGFNVSAVALQLGPK--DASRQVYICSNNIQARQOVIELARQNLFIPIDLGSL 196	
DB	119	GAKVKAFTNLPAAVLAADPDKGTGSRVLFISGNHSDANRQVAELISSIGFAFVDLGTLA 178	
QY	197	SAREIENPLRLFTLWRGPVVAISL 222	
DB	179	ASGPIQOF-----GRPLVALNL 195	

RESULT 2

AC2560
hypothetical protein alr8074 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120ga
C/Species: Nostoc sp. PCC 7120
A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C/Accession: AC2560
R/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A/Reference number: AB1807; MUID:21595285; PMID:11759840
A/Accession: AC2560
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-211 <KUR>
A/Cross-references: GB:AP003603; PIDN:BAB77404.1; PID:g17134847; GSPDB:GN00182
A/Experimental source: strain PCC 7120
C/Genetics:
A/Gene: alr8074
A/Genome: plasmid

```
Query Match      6.8%; Score 159.5; DB 2; Length 211;
Best Local Similarity 24.9%; Pred. No. 2.4e-05;
Matches 51; Conservative 43; Mismatches 70; Indels 41; Gaps 8;
```

[illegible]

```

RESULT 3
T50571
probable oxidoreductase [imported] - Streptomyces coelicolor
C:/Species: Streptomyces coelicolor
C:/Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 28-Jul-2000

```

R;Redenbach, M., Kieser, H.M.; Denapalte, D.; Eichner, A.; Culm, J.; Kinashi, H.; Hopmann, M. *Microbiol.* 21, 77-96, 1996
A;Title: A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb S. *flexneri* 304 strain
A;Reference number: Z20556; MUID:97000351; PMID:8843436
A;Accession: T50571
A;Status: Preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-239 <RED>
A;Cross-references: EMBL:AL133220; PIDN:CAB61708.1
A;Experimental source: strain A3(2)
C;Genetics:
A;Note: SCC75A.08c
C;Superfamily: conserved hypothetical protein MJ1501

Query Match	6.7%;	Score 156.5;	DB 2;	Length 239;
Best Local Similarity	27.9%;	Pred. No. 5e-05;		
Matches 61;	Conservative 41;	Mismatches 72;	Indels 45;	Gaps 10;

DQ
25 IKDARKYIVGI - GSGDFAKSLTIRLIRCGHYVIGSR --- NPKFASEFFPHVDYTHHE 80
 : : : : : : : : : : : : : : : : :
Db 23 LPDVSGELVGLGTGPQGKGLAYRLAKAGQKIVGSRRAERAAAAABEIGHGVGADNA 82
 : : : : : : : : : : : : : : : : :

QY 81 DALTKNIIIFVAIHRE - HYTSIWMDLRHLVSKILIDVSNNMRINO ----- YPE -- SNAE 131
 : : : : : : : : : : : : : : : : :
Db 83 ETARRSDVIIVAVPMDGHGKTLESIRAEISGKIYVDCVNPLGFDKGAYALKPESGAEE 142

```
QY      132 YLASLPEDSLIVKGFNVSAWALQ-----LGPKDASRQVYICSNNTGARQQV    178
           |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      143 QAAALLPDSRVAAAFFHHLTAVLLQDP EIDEIDTDVMVLGEERADVEI-----YQA-----    192
           |||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      179 IELARQLNFI P----IDLGIS SARIEINLPRLFTIMR    213
           |||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      193 --LAGR---IPGMRGVFAGRLRNAHQVESLVANLISVNR    226
```

RESULT 4
H69400

C:\Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 29-Sep-1999
C:\Accession: H69400

R./Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997

A.; Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Attiach, P.; Kaine, B.P.; Sykes, S. Smith, H.O.; Woese, C.R.; Venter, J.C.

A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon
A;Reference number: A69250; MUID:98049343; PMID:9389475

A;Accession: H69400

A; Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Residues: 1-213 <KLE>
A;Cross-references: GB:AE001021; NID:g26689344; PIDN:AAB90038.1; PID:g2649337
C;Superfamily: conserved hypothetical protein MJ1501

Query Match	6.3%;	Score 147;	DB 2;	Length 213;
Best Local Similarity	27.5%;	Pred. NO. 0.00024;		
Matches	56;	Conservative	44;	Mismatches 68;
			Indels	36;
			Gaps	11;

```
OY      37 GSQDFAKSLTRLRLRCGYHVIIGSRN---PKFASEFFPHVDVT---HHEDALTKTNI   88
       ||:::||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      8 GTGNLGEGLALRWGKLGYEIIIVGSRKLEKAETLASDYLLKKVGDASIIGRNEDAETCD-  66
```

```
QY      89 IFVAIRKHYTSLMD-----LRHLLVGKILID-----VSNMNRINQYPE--SNAE 131
       ||| : | : ||| : |||
Db     67 --VAV----FTIPWEFAFDIAEMLKRLAGKXVISPLVPMKKVGDNF-VYYRPEEGSAAE 119
```

```
QY      191 DLGLSSAREIENL-PLRLFTLWR 213
      179 DAGLSNAHLVESTPLILNVMKR 202
Db
```

RESULT 5
D69361

conserved hypothetical protein AF0892 - *Archaeoglobus fulgidus*

C;Species: Archaeoglobus fulgidus

C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 29-Sep-1999
C;Accession: D69361

R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodet, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A./Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artlach, P.; Kaine, B.P.; Sykes, S. Smith, H.O.; Woese, C.R.; Venter, J.C.

A: Reference number: A59250; MTID: 98049343; PMT: 9389475

A: Accession: D69361

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A; molecule type: DNA

A;Residues: 1-212 <KLE>

A; Cross-references: GB:AE001042, NID:g22689365, PIDN:AAB90348.1, PID:g2264970
C; Superfamily: conserved hypothetical protein MJ1501

[illegible][illegible]

A/Reference number: A64300; MUID:96337999; PMID:8688087
A/Accession: D64487
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-223 <BUT>
A/Cross-references: GB:U67591; GB:L77117; NID:g2826422; PIDN:AAB99514.1; PID:g1500389; T
C/Genetics:
A/Map position: REV1473617-1472946
C/Superfamily: conserved hypothetical protein MJ1501

	Query Match	5.7%	Score 133;	DB 2;	Length 223;	
	Best Local Similarity	21.9%;	Pred. No. 0.0033;			
	Matches	46;	Conservative	52;	Mismatches	78;
					Indels	34;
					Gaps	7
QY	37	GS	GDFAKSLTIRLRCGYHVIVIGSRNPKFASFFPHVDV-----THEDALT	84		
		:	: :	:	:	:
		:	:	:	:	:
Db	8	GTGDQGFGLATRLAK-NNKIIISRKKEKAEMAKKAKEILKQGIHADITGLENKDPAK	66			
QY	85	KTNIIIFVAIHREH-YTSLMDLRHLVLVGKILIDV-----SNMNRINQYPE-SNAEYL	133			
		:	:	:	:	:
		:	:	:	:	:
Db	67	EGDVVILSLPYEYTLSTIKQLKEELKGKIIVSIGVPILATAGDKPTRLLFPDGSAVMV	126			
QY	134	ASLFPPDSLTVKGFNVVSAMALQLGPKDASRQVYICSSNNIQARQCVIELARQLNFI-PIDL	192			
		:	:	:	:	:
		:	:	:	:	:
Db	127	QNVLKESKVSAFQNVCHAVLEDDLNPVDCDILVCGNDEEAKKVVIDLANQIDGVRAIDC	186			
QY	193	GSLSAREIENLPRLFTLMRGPVVAISL	222			
		:	:	:	:	:
		:	:	:	:	:
Db	187	GNLEKSRIIEAIT-----PLILGLNI	207			

RESULT 8
F86826
hypothetical protein yqfE [imported] - *Lactococcus lactis* subsp. *lactis* (strain IL1403)
C/Species: *Lactococcus lactis* subsp. *lactis*
C/Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C/Accession: F86826
R/Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A/Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ss
A/Reference number: A86625; MUID:21235186; PMID:11337471
A/Accession: F86826
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-191 <STO>
A/Cross-references: GB:AE005176; PID:g12724621; PIDN:AAK05712.1; GSPDB:GN00146
A/Experimental source: strain IL1403
C/Genetics:
A/Gene: yqfE

	Query Match	5.1%;	Score 121;	DB 2;	Length 191;	
	Best Local Similarity	22.2%;	Pred. No. 0.024;			
	Matches	47;	Conservative	33;	Mismatches	74;
					Indels	58;
					Gaps	7
QY	32	TWGVIGSGDFAKSLTIRLRGCIYHVIVIGSRNPKFASFEFFPHVVDYTHHEDALTKTN----	87			
Db	3	TISIFGKMGKAI-----	GDNFSSSVKNKYITLSSSKTELG EI	42		
QY	88	-----IIFVAIHREHYTSMDRLHLVGKILIDVSNNMRINQYE-----SNAEYL	133			
Db	43	VVLAVPYVALAGIIQEYSTD-----LQGKIIDIITNPVDTTFPSLLVPSDTSAALI	95			
QY	134	ASLEPDSLIVKGFNVSAMALOLGPKDASRO---VYICSNNICARQOVIELARQLNFIP	169			
Db	96	AKQLPNSMIVKAFN--TTESDYLATKKVANEHQTVVLLASDSQEAKEITIKALENSGLSL	153			
QY	190	IDLGLSSAREIENTPLRLFTL-----WRG	214			
Db	154	LDAGSLKRARELEAIGFLQTITLAASEKISWDG	185			

RESULT 9
A69131

conserved hypothetical protein MTH248 - Methanobacterium thermoautotrophicum (strain Del
C/Species: Methanobacterium thermoautotrophicum
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 29-Sep-1999
C/Accession: A69131
R/Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
kl, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A/Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A/Reference number: A69000; MUID:98037514; PMID:9371463
A/Accession: A69131
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-232 <MTH>
A/Cross-references: GB:AE000811; GB:AE000666; NID:G2621287; PIDN:AAB84754.1; PID:G262129
A/Experimental source: strain Delta H
C/Genetics:
A/Gene: MTH248
A/Start codon: TTG
C/Superfamily: conserved hypothetical protein MJ1501

Query Match	5.1%;	Score 120.5;	DB 2;	Length 232;
Best Local Similarity	23.8%;	Pred. No. 0.034;		
Matches 49;	Conservative 42;	Mismatches 90;	Indels 25;	Gaps 6;

```

QY      30 KVTWGVIGSGDFAKSLTIRLCGYHWVIGSRNPKASEFPFHVVDVTHHE----- 80
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db      8 KIAV-IGTGDGGLALRAVAGEEVIIGSRDAEKASKASKVLEIAGRDDISYEGATN 66
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

```

QY      81 -DALTKNIIFVAIH-REHYTSLMDLRHLLVGKILIDVS-----NNMRINQYPSN 129
      || ::::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::
Db      67 PDAASADVVLTVPLOQMWTLASIRDPQVRKVLIDATVPIDSCIGSAVRITDMEGS 126

```

```

QY      130 AEYLASLF---PDSLIVKGFNVVSAMALQLGPKDASRQVYICSNNIQARQCVITELAROLN 186
          |  :  |  :  |  :  |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      127 AAERARAFIRKQGT RVA AAFNNISASALLEVSEPVDCDLVASDHRDALEVA AELAEKID 186

```

```

QY      187 FI-PIDGLSSAREIENLPRLFTL 211
      : | : | : | : | : |
DB      187 GYRAIECGGLENARIIEKITPPLINL 212

```

RESULT 10
T00121

hypothetical protein 8 - *Leptospira interrogans*
C;Species: *Leptospira interrogans*

C:/Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
C:/Accession: T00121
R; Takahashi, Y.; Akase, K.; Hirano, H.; Fukunaga, M.

Gene 215, 37-45, 1998

A;Title: Physical and genetic maps of the *Leptospira interrogans* serovar icterohaemorrhagiae

A;Reference number: Z14115; MUID:98332717; PMID:96660700

A/Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-216 <TAK>

A; Experimental source: strain Ictero No.1; substrain icterohaemorrhagiae

Query Match 5.1%; Score 119.5; DB 2; Length 216;

DESeq Local Similarity 23.78; Acc: 0.957;
 Matches 47; Conservative 37; Mismatches 91; Indels 23; Gaps 6;

QY 33 VGVIGSGDFAKSLTIRLCGYHVVIGSRNPKFASEFFPHV--VDVTHHEDALTKNII 89

Db 6 I G I L S G I V G Q T L A N G F L K Y G A E V K I G T R D F G K L K D W L A K A G A G A S I G S F S E A N F G E I I

QY 90 FVAIHREHYTSLWDRLH--LVGKILIDVSNMR-----INQYPESNAEYLA SL 136

Db 66 VLCSKGSVASEVLTLSGIDSLNGKTIIDTNPISSEIPPQNGVLNFFTSYNESLMEKLQKQ 125

QY 137 FPD SLIVKFNVSAMALQLGP--KDASRQVYICSNNIQARQVIELARQLNFIPI DLGS 194

```
Db      126 A P K A N F Y K C F S S V G S - G L M V N P Q L K G E K P S M F I C G N D S S K O I K E I L D T F G W D T E D M G K   184
           | : | | : | | : | | : | : : : | :
Qy      195 L S S A R E I E N L P I R L F T L W    212
           : : | | | : |||
Db      185 V E A A R A T E P L C I --- L W    198
```

RESULT 11
G82642

C,Species: Xylella fastidiosa
C,Date: 18-Aug-2000 #sequence 20-Aug-2000 #text_change 20-Aug-2000

R;aronymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
Nature 406, 151-157, 2000

A: Note: for a complete list of authors see reference number A59328 below
A: Reference number: A82515; MUID:20365717; PMID:10910347

A:Accession: G02072
A:Status: preliminary
A:Molecule type: DNA

A:Residue: 1-242 <SIM>
A:Cross-References: GB:AE003997; GB:AE003849; NID:g9106805; PIDN:AAF84546.1; GSPDB:GN001
A:Experimental source: strain 9a5c

R/Simpson, A.J.G.; Keinacny, F.C.; Alruad, F.; Andru, F.A.; Abenculo, M.; Alvaranga, N.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H. Neto, F.; Docena, C.; El-Dorrry, H.; Facincani, A.P.; Ferreira, A.J.S.;

submitted to GenBank, June 2000
A.; Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohn
I.D.; Truett, M.T.; Brown, E.L.; Vitalina, J.P.; Krieger, J.E.; Kuramae, E.F.; Lajtha
K.

chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.
A.; Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
Muniz, D.Z.; Oliveira, P.C.; Palmieri, D.Z.; Pereira, M.A.; de Oliveira, R.C.

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawashiro, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silva, A.; Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silva, A.; Authors:

A;Reference number: A59328
A;Contents: annotation

A;Gene: XF1737

Query Match	5.0%;	Score 116.5;	DB 2;	Length 242;
Best Local Similarity	24.0%;	Pred. No. 0.075;		
Matches 46;	Conservative 40;	Mismatches 85;	Indels 21;	Gaps 7;

QY 28 ARKVTWGVIGSGDFAKSLTILRLRCGYHWVIGSRNP----KFASEFFPHVDVTHHEDAL 83
| : |||| : :: : |
Db 30 AAPMRIGIVAGSLGGTYGRLLMWKAGHEVMFSSRNPDKLEAMARELEPR-ASVGQPLAAT 88

```
QY      84 TKTNIIFVAIHREHYTSLM-DLRHLLVGKILIDVSNMRINQ---YPESN----AETLAS   135
       :::::  : ||| ::::: | : :
Db     89 EFGTVLLLVPFEEALPQVGRDLRSAYRGKIVLDSTNPWGASSADVYREARELGVAGQTIVK   148
```

```

QY      136 LFPDSLIVKGENVVSAMALQIGPKDSRQ-----VYICSNIIQARQVIELARQLNFLEPI 190
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      149 YMPGARLVRAPSAVDAITVE---TSASRRCGRIGMPLASDDAEAKVAEGLVRDAGCDEY 205

```

```
QY      191 DLGLSSAREIE 202
          :|::|:
Db      206 IVGNLAAAASFQ 217
```

RESULT 12
AB3182

conserved hypothetical protein Atu5183 [imported] - Agrobacterium tumefaciens (strain C58)
C/Species: Agrobacterium tumefaciens

C/Accession: AB3182
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Cher, L.; Wood, G.E.; Chen, Y.; Woo, I.

Karp, P.; Romero, P.; Zhang, S. Science 294, 2317-2323, 2001

[illegible]

	Query Match	4.7%; Score 111.5; DB 1;	Length 695;
	Best Local Similarity	17.0%; Pred. No. 0.71;	
	Matches	84; Conservative	75; Mismatches 151; Indels 183; Gaps 16;
OY	79 HEDALTKNTIIFVAIHREHYTSLMDIRHLLVG-----KILIDVSNNMR	121	
Dd	98 HEIRIEKANNL-LYNPEAFQNLPLRYLLISNLTGIKLIPDVHKHSFQKVLLDIOENIN	156	
OY	122 INQPESNAEYLASLPDSLIVKGENVVASAWALOGLPKDASROYICSNNIQARQQVIEL	181	
Dd	157 IH-----TIERNSEFGVLSPFSVITWL---NKNGIQEIHNCA-----	189	
OY	182 ARQLNFPIPDGLSSAREIENTPRLFTLMRGPVVAISLATFFFLYSFVRDIVHPYAR	241	
Dd	190 ---FNQTGLELNLSDNNNLEELPNDDVFHGASGPVILLISRTRISHSLPSYLENLKKLR	246	
OY	242 NQOSDFYKPPIETIVNKTLPIVAATILLSLVY-----	271	
Dd	247 RSTYNLKKLPP----SLEKLVALMEASLYTPSHCCAFANWRROISELHPICKNSILRQEY	301	
OY	272 -----LAGLLAAAYQLYYGTKYRRF-----PWLETWLOCRXQLG	306	
Dd	302 DYMOTGRQRSSLAEADNESSYSRGFDMTYAEFDYDLCENEVDVTCSFKPDFAFPNCEDIIG	361	
OY	307 ----LTSFFPAWHVA-----YSLCLPMRRSERLYFLNMAYQQOVHANIEWS	348	
Dd	362 YNILRVLIWFISI LAITGNII VLVTLTTSQYKLTVP----RFLMGNLAFADE-----	408	
OY	349 WNEEEVWRILEMYISFGIMSLGLSLAVTSIPSUSNALNWRE-----SFSIQSTLG	399	
Dd	409 -----LCIGIYLLLIASVDITHKSQYHNVAIDMQTGAGCDAGFFTASFASELS	456	
OY	400 ---YVALLISTFH-----VLITYGWKRAFEERYR.FYTPTPNFV-IAL	436	
Dd	457 VYTLTAITLERWHTITHAMOLDCRKVVHRHAASVMWMGMIFAFAAALFPIFGISSYMKVSI	516	
OY	437 VLPSIVIIDL LQL 449		
Dd	517 CLPMDIDSPLSQL 529		

RESULT 14
559681
probable membrane protein YPL012w - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: hypothetical protein LPB5w; hypothetical protein YPB132.01
C:Species: Saccharomycetes cerevisiaae
C>Date: 13-Jan-1996 #sequence_revision_01-Mar-1996 #text_change 19-Apr-2002
C/Accession: S59681; S52519
R:Hail, J.; Ahmed, A.; Bussey, H.; Fortin, N.; Friesen, J.D.; Storms, R.K.; Vo, D.H.; Wa submitted to the EMBL Data Library, August 1995
A>Description: The sequence of *Saccharomyces cerevisiae* chromosome XVI left arm.
A:Reference number: S59677
 A:/Accession: S59681
 A/Molecule type: DNA
 A;/Residues: 1-1228 <HAL>
A;/Cross-references: EMBL:U33335; NID:g965076; PID:g965081; MIPS:YPL012w
 R/Badcock, K.; Churcher, C.
submitted to the EMBL Data Library, February 1995
 A:/Reference number: S52519
 A/Accession: S52519
 A/Molecule type: DNA
 A;/Residues: 220-1228 <BAD>
A;/Cross-references: EMBL:Z48483; NID:g683777; PID:g683778; MIPS:YPL012w
 A;/Experimental source: strain AB972
 C/Genetics:
 A/Gene: SGD:RRP12
 A;/Cross-references: SGD:S0005933
 A/Map position: 16L
 C/Keywords: transmembrane protein
F:125-141/Domain: transmembrane #status predicted <TM1>
F:478-494/Domain: transmembrane #status predicted <TM2>
F:720-736/Domain: transmembrane #status predicted <TM3>


```

; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US01/06520
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06666
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10
; SEQ ID NO 10
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-09-888-257A-10
```

```

Query Match      100.0%; Score 2351; DB 10; Length 454;
Best Local Similarity 100.0%; Pred. No. 6.1e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      1 MESISMGMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVVIGS 60
      |||
      1 MESISMGMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVVIGS 60
DB
QY      61 RNPKFASEFFPHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLLVCKILIDVSNM 120
      |||
      61 RNPKFASEFFPHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLLVCKILIDVSNM 120
DB
QY      121 RINQYPESNAEYLAFLPDSLIVKGFNVSAWALQLGPKDASRQVYICSNNIQARQVIE 180
      |||
      121 RINQYPESNAEYLAFLPDSLIVKGFNVSAWALQLGPKDASRQVYICSNNIQARQVIE 180
DB
QY      181 LARQLNFIPIDLGSLSSAREIENLPLRLFTLWGFVVAISLATFFFLYSFVRDVIHPYA 240
      |||
      181 LARQLNFIPIDLGSLSSAREIENLPLRLFTLWGFVVAISLATFFFLYSFVRDVIHPYA 240
DB
QY      241 RNQOSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTXYRRPFWLETWLQ 300
      |||
      241 RNQOSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTXYRRPFWLETWLQ 300
DB
QY      301 CRKQLGLLSFFPAMVHYAAYSLCLPMRSEERYLFNMAYQQOVHANINENSWNEEVEWRIEMY 360
      |||
      301 CRKQLGLLSFFPAMVHYAAYSLCLPMRSEERYLFNMAYQQOVHANINENSWNEEVEWRIEMY 360
DB
QY      361 ISFGIMSLGLSLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHYLIYGMKRAFE 420
      |||
      361 ISFGIMSLGLSLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHYLIYGMKRAFE 420
DB
QY      421 BEYYRFTYTPNPFVLAIVLPSIVILDLQLCRYPD 454
      |||
      421 BEYYRFTYTPNPFVLAIVLPSIVILDLQLCRYPD 454
DB
```

RESULT 2

```

US-10-455-822-3
; Sequence 3, Application US/10455822
; Publication No. US20040048798A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Farris, Mary
; TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer
; FILE REFERENCE: 51158-20016.24
; CURRENT APPLICATION NUMBER: US/10/455,822
; PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US60/370,387
; PRIOR FILING DATE: 2002-04-05
```

```

; PRIOR APPLICATION NUMBER: US60/317,840
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: US60/296,656
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US60/091,183
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US60/087,520
; PRIOR FILING DATE: 1998-06-01
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Homo sapien
; US-10-455-822-3
```

```

Query Match      100.0%; Score 2351; DB 12; Length 454;
Best Local Similarity 100.0%; Pred. No. 6.1e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      1 MESISMGMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVVIGS 60
      |||
      1 MESISMGMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVVIGS 60
DB
QY      61 RNPKFASEFFPHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLLVCKILIDVSNM 120
      |||
      61 RNPKFASEFFPHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLLVCKILIDVSNM 120
DB
QY      121 RINQYPESNAEYLAFLPDSLIVKGFNVSAWALQLGPKDASRQVYICSNNIQARQVIE 180
      |||
      121 RINQYPESNAEYLAFLPDSLIVKGFNVSAWALQLGPKDASRQVYICSNNIQARQVIE 180
DB
QY      181 LARQLNFIPIDLGSLSSAREIENLPLRLFTLWGFVVAISLATFFFLYSFVRDVIHPYA 240
      |||
      181 LARQLNFIPIDLGSLSSAREIENLPLRLFTLWGFVVAISLATFFFLYSFVRDVIHPYA 240
DB
QY      241 RNQOSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTXYRRPFWLETWLQ 300
      |||
      241 RNQOSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTXYRRPFWLETWLQ 300
DB
QY      301 CRKQLGLLSFFPAMVHYAAYSLCLPMRSEERYLFNMAYQQOVHANINENSWNEEVEWRIEMY 360
      |||
      301 CRKQLGLLSFFPAMVHYAAYSLCLPMRSEERYLFNMAYQQOVHANINENSWNEEVEWRIEMY 360
DB
QY      361 ISFGIMSLGLSLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHYLIYGMKRAFE 420
      |||
      361 ISFGIMSLGLSLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHYLIYGMKRAFE 420
DB
QY      421 BEYYRFTYTPNPFVLAIVLPSIVILDLQLCRYPD 454
      |||
      421 BEYYRFTYTPNPFVLAIVLPSIVILDLQLCRYPD 454
DB
```

RESULT 3

```

US-10-455-822-7
; Sequence 7, Application US/10455822
; Publication No. US20040048798A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Farris, Mary
; TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer
; FILE REFERENCE: 51158-20016.24
; CURRENT APPLICATION NUMBER: US/10/455,822
; PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US60/370,387
; PRIOR FILING DATE: 2002-04-05
```

```

; PRIOR APPLICATION NUMBER: US60/317,840
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: US60/296,656
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US60/091,183
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US60/087,520
; PRIOR FILING DATE: 1998-06-01
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Homo sapien
; US-10-455-822-7

```

```

Query Match      100.0%; Score 2351; DB 12; Length 454;
Best Local Similarity 100.0%; Pred. No. 6.1e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVIGS 60
DB 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVIGS 60
QY 61 RNPKFASEFPFHVVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVGKILLIDVSNM 120
DB 61 RNPKFASEFPFHVVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVGKILLIDVSNM 120
QY 121 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQIGPKDASROVYICSNIQARQVIE 180
DB 121 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQIGPKDASROVYICSNIQARQVIE 180
QY 121 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQIGPKDASROVYICSNIQARQVIE 180
DB 121 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQIGPKDASROVYICSNIQARQVIE 180
QY 181 LARQJNFIPIDGLSSAREIENLPLRLFTLWPGPVVAISLATFFFLYSFVVDVIHPYA 240
DB 181 LARQJNFIPIDGLSSAREIENLPLRLFTLWPGPVVAISLATFFFLYSFVVDVIHPYA 240
QY 241 RNOQSDFYKIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYGTKYRRFPWLETWLQ 300
DB 241 RNOQSDFYKIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYGTKYRRFPWLETWLQ 300
QY 301 CRKQGLLSFFFAMVHVAYSCLPMRSEERYLFNMAVQOVHANIENSWNEEYWRIMY 360
DB 301 CRKQGLLSFFFAMVHVAYSCLPMRSEERYLFNMAVQOVHANIENSWNEEYWRIMY 360
QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRREFSIQSTLGYVALLISTFHVLIYGMKRAFE 420
DB 361 ISFGIMSLGLSLAVTSIPSVSNALNWRREFSIQSTLGYVALLISTFHVLIYGMKRAFE 420
QY 421 EBYRYFTPPNFVLAIVLPISIVIDLQLCRYPD 454
DB 421 EBYRYFTPPNFVLAIVLPISIVIDLQLCRYPD 454

```

```

RESULT 4
US-10-455-822-9
; Sequence 9, Application US/10455822
; Publication No. US20040048798A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya
; APPLICANT: Chailita-Eid, Pia M.
; APPLICANT: Paris, Mary
; TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer
; FILE REFERENCE: 51158-20016.24
; CURRENT APPLICATION NUMBER: US/10/455,822
; PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US60/370,387
; PRIOR FILING DATE: 2002-04-05

```

```

; PRIOR APPLICATION NUMBER: US60/317,840
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: US60/296,656
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US60/091,183
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US60/087,520
; PRIOR FILING DATE: 1998-06-01
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Homo sapien
; US-10-455-822-9

```

```

Query Match      100.0%; Score 2351; DB 12; Length 454;
Best Local Similarity 100.0%; Pred. No. 6.1e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVIGS 60
DB 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVIGS 60
QY 61 RNPKFASEFPFHVVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVGKILLIDVSNM 120
DB 61 RNPKFASEFPFHVVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVGKILLIDVSNM 120
QY 121 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQIGPKDASROVYICSNIQARQVIE 180
DB 121 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQIGPKDASROVYICSNIQARQVIE 180
QY 121 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQIGPKDASROVYICSNIQARQVIE 180
DB 121 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQIGPKDASROVYICSNIQARQVIE 180
QY 181 LARQJNFIPIDGLSSAREIENLPLRLFTLWPGPVVAISLATFFFLYSFVVDVIHPYA 240
DB 181 LARQJNFIPIDGLSSAREIENLPLRLFTLWPGPVVAISLATFFFLYSFVVDVIHPYA 240
QY 241 RNOQSDFYKIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYGTKYRRFPWLETWLQ 300
DB 241 RNOQSDFYKIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYGTKYRRFPWLETWLQ 300
QY 301 CRKQGLLSFFFAMVHVAYSCLPMRSEERYLFNMAVQOVHANIENSWNEEYWRIMY 360
DB 301 CRKQGLLSFFFAMVHVAYSCLPMRSEERYLFNMAVQOVHANIENSWNEEYWRIMY 360
QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRREFSIQSTLGYVALLISTFHVLIYGMKRAFE 420
DB 361 ISFGIMSLGLSLAVTSIPSVSNALNWRREFSIQSTLGYVALLISTFHVLIYGMKRAFE 420
QY 421 EBYRYFTPPNFVLAIVLPISIVIDLQLCRYPD 454
DB 421 EBYRYFTPPNFVLAIVLPISIVIDLQLCRYPD 454

```

```

RESULT 5
US-10-455-822-19
; Sequence 19, Application US/10455822
; Publication No. US20040048798A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya
; APPLICANT: Chailita-Eid, Pia M.
; APPLICANT: Paris, Mary
; TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer
; FILE REFERENCE: 51158-20016.24
; CURRENT APPLICATION NUMBER: US/10/455,822
; PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US60/370,387
; PRIOR FILING DATE: 2002-04-05

```


PRIOR APPLICATION NUMBER: US60/317,840
PRIOR FILING DATE: 2001-09-06
PRIOR APPLICATION NUMBER: US60/296,656
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US60/091,183
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US60/087,520
PRIOR FILING DATE: 1998-06-01
NUMBER OF SEQ ID NOS: 197
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 19
LENGTH: 454
TYPE: PRT
ORGANISM: Homo sapien
US-10-455-822-19

Query Match 100.0%; Score 2351; DB 12; Length 454;
Best Local Similarity 100.0%; Pred. No. 6.1e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPENGINGIKDARKVTGVIIGSGDFAKSLTIRLCGYHVVIGS 60
DB 1 MESISMGSPKSLSETCLPENGINGIKDARKVTGVIIGSGDFAKSLTIRLCGYHVVIGS 60
QY 61 RNPKEASEFFPHVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVGKILLIDVSNM 120
DB 61 RNPKEASEFFPHVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVGKILLIDVSNM 120
QY 121 RINQYPSNAEYLASLFPDSLIVKGFNVVSAMALQGPKDASROVYICSNNIQARQVIE 180
DB 121 RINQYPSNAEYLASLFPDSLIVKGFNVVSAMALQGPKDASROVYICSNNIQARQVIE 180
QY 181 LARQINFIPIDLGSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVDVIHPYA 240
DB 181 LARQINFIPIDLGSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVDVIHPYA 240
QY 241 RNOQSDFYKIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYGTXYRRFPWLETWLQ 300
DB 241 RNOQSDFYKIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYGTXYRRFPWLETWLQ 300
QY 301 CRKQGLLSFFFAMVHVAAYSLCLPMRSEERYLFLNMAYQOVHANIEINSWNEEYWRIMY 360
DB 301 CRKQGLLSFFFAMVHVAAYSLCLPMRSEERYLFLNMAYQOVHANIEINSWNEEYWRIMY 360
QY 361 ISFGIMSGLLSLAVTSIPSVSNALNWRRESFIQSTLGYVALLISTFHVLIYGMKRAFE 420
DB 361 ISFGIMSGLLSLAVTSIPSVSNALNWRRESFIQSTLGYVALLISTFHVLIYGMKRAFE 420
QY 421 EYYRFTYPPNFVLALVLPISIVLIDLQLCRYPD 454
DB 421 EYYRFTYPPNFVLALVLPISIVLIDLQLCRYPD 454

RESULT 6
US-10-455-822-21
Sequence 21, Application US/10455822
Publication No. US20040048798A1
GENERAL INFORMATION:
APPLICANT: Agensys, Inc.
APPLICANT: Raitano, Arthur B.
APPLICANT: Ge, Wangmao
APPLICANT: Jakobovits, Aya
APPLICANT: Chailita-Eid, Pia M.
APPLICANT: Farris, Mary
TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer
FILE REFERENCE: 51158-20016.24
CURRENT APPLICATION NUMBER: US/10/455,822
CURRENT FILING DATE: 2003-06-04
PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
PRIOR FILING DATE: 2002-12-20
PRIOR APPLICATION NUMBER: US60/370,387
PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
PRIOR FILING DATE: 2001-09-06
PRIOR APPLICATION NUMBER: US60/296,656
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US60/091,183
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US60/087,520
PRIOR FILING DATE: 1998-06-01
NUMBER OF SEQ ID NOS: 197
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 21
LENGTH: 454
TYPE: PRT
ORGANISM: Homo sapien
US-10-455-822-21

Query Match 100.0%; Score 2351; DB 12; Length 454;
Best Local Similarity 100.0%; Pred. No. 6.1e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPENGINGIKDARKVTGVIIGSGDFAKSLTIRLCGYHVVIGS 60
DB 1 MESISMGSPKSLSETCLPENGINGIKDARKVTGVIIGSGDFAKSLTIRLCGYHVVIGS 60
QY 61 RNPKEASEFFPHVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVGKILLIDVSNM 120
DB 61 RNPKEASEFFPHVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVGKILLIDVSNM 120
QY 121 RINQYPSNAEYLASLFPDSLIVKGFNVVSAMALQGPKDASROVYICSNNIQARQVIE 180
DB 121 RINQYPSNAEYLASLFPDSLIVKGFNVVSAMALQGPKDASROVYICSNNIQARQVIE 180
QY 181 LARQINFIPIDLGSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVDVIHPYA 240
DB 181 LARQINFIPIDLGSLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVDVIHPYA 240
QY 241 RNOQSDFYKIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYGTXYRRFPWLETWLQ 300
DB 241 RNOQSDFYKIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYGTXYRRFPWLETWLQ 300
QY 301 CRKQGLLSFFFAMVHVAAYSLCLPMRSEERYLFLNMAYQOVHANIEINSWNEEYWRIMY 360
DB 301 CRKQGLLSFFFAMVHVAAYSLCLPMRSEERYLFLNMAYQOVHANIEINSWNEEYWRIMY 360
QY 361 ISFGIMSGLLSLAVTSIPSVSNALNWRRESFIQSTLGYVALLISTFHVLIYGMKRAFE 420
DB 361 ISFGIMSGLLSLAVTSIPSVSNALNWRRESFIQSTLGYVALLISTFHVLIYGMKRAFE 420
QY 421 EYYRFTYPPNFVLALVLPISIVLIDLQLCRYPD 454
DB 421 EYYRFTYPPNFVLALVLPISIVLIDLQLCRYPD 454

RESULT 7
US-10-455-822-23
Sequence 23, Application US/10455822
Publication No. US20040048798A1
GENERAL INFORMATION:
APPLICANT: Agensys, Inc.
APPLICANT: Raitano, Arthur B.
APPLICANT: Ge, Wangmao
APPLICANT: Jakobovits, Aya
APPLICANT: Chailita-Eid, Pia M.
APPLICANT: Farris, Mary
TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer
FILE REFERENCE: 51158-20016.24
CURRENT APPLICATION NUMBER: US/10/455,822
CURRENT FILING DATE: 2003-06-04
PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
PRIOR FILING DATE: 2002-12-20
PRIOR APPLICATION NUMBER: US60/370,387
PRIOR FILING DATE: 2002-04-05

```

; PRIOR APPLICATION NUMBER: US60/317,840
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: US60/296,656
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US60/091,183
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US60/087,520
; PRIOR FILING DATE: 1998-06-01
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-455-822-23
```

```

Query Match      100.0%; Score 2351; DB 12; Length 454;
Best Local Similarity 100.0%; Pred. No. 6.1e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      1 MESISMGSPKSLSETCLPENGINGIKDARKVTGVTGSGDFAKSLTIRLRGCVHYVIGS 60
         |||
Db      1 MESISMGSPKSLSETCLPENGINGIKDARKVTGVTGSGDFAKSLTIRLRGCVHYVIGS 60
QY      61 RNPKEASEFFPHVVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
         |||
Db      61 RNPKEASEFFPHVVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
QY      121 RINQYPSNAEYLASLFPDSLIVKGFNVSAWALQGPKDASROYVICSNNIQARQVIE 180
         |||
Db      121 RINQYPSNAEYLASLFPDSLIVKGFNVSAWALQGPKDASROYVICSNNIQARQVIE 180
QY      121 RINQYPSNAEYLASLFPDSLIVKGFNVSAWALQGPKDASROYVICSNNIQARQVIE 180
         |||
Db      121 RINQYPSNAEYLASLFPDSLIVKGFNVSAWALQGPKDASROYVICSNNIQARQVIE 180
QY      181 LARQLNFIPIIDLGLSSAREIENLPLRLFTLWRGPVVVAISLATFFFLYSFVRDVIHPYA 240
         |||
Db      181 LARQLNFIPIIDLGLSSAREIENLPLRLFTLWRGPVVVAISLATFFFLYSFVRDVIHPYA 240
QY      181 LARQLNFIPIIDLGLSSAREIENLPLRLFTLWRGPVVVAISLATFFFLYSFVRDVIHPYA 240
         |||
Db      181 LARQLNFIPIIDLGLSSAREIENLPLRLFTLWRGPVVVAISLATFFFLYSFVRDVIHPYA 240
QY      241 RNOQSDFYKPIPIEIVNKTLPPIVATLLSLVYLAGLLAAAYQLYYGTYRRFPFWLETWLO 300
         |||
Db      241 RNOQSDFYKPIPIEIVNKTLPPIVATLLSLVYLAGLLAAAYQLYYGTYRRFPFWLETWLO 300
QY      241 RNOQSDFYKPIPIEIVNKTLPPIVATLLSLVYLAGLLAAAYQLYYGTYRRFPFWLETWLO 300
         |||
Db      241 RNOQSDFYKPIPIEIVNKTLPPIVATLLSLVYLAGLLAAAYQLYYGTYRRFPFWLETWLO 300
QY      301 CRKQLGLISFFPAMVHVAYSLCLPMRSEERYFLNMAYQOVHANINENSWNEEVRRIEMY 360
         |||
Db      301 CRKQLGLISFFPAMVHVAYSLCLPMRSEERYFLNMAYQOVHANINENSWNEEVRRIEMY 360
QY      301 CRKQLGLISFFPAMVHVAYSLCLPMRSEERYFLNMAYQOVHANINENSWNEEVRRIEMY 360
         |||
Db      301 CRKQLGLISFFPAMVHVAYSLCLPMRSEERYFLNMAYQOVHANINENSWNEEVRRIEMY 360
QY      361 ISFGIMSLGLLSLAVTSIPSVSNALNWRREFSIQSTLGVAALLISTFHVLIYGWKRAFE 420
         |||
Db      361 ISFGIMSLGLLSLAVTSIPSVSNALNWRREFSIQSTLGVAALLISTFHVLIYGWKRAFE 420
QY      361 ISFGIMSLGLLSLAVTSIPSVSNALNWRREFSIQSTLGVAALLISTFHVLIYGWKRAFE 420
         |||
Db      361 ISFGIMSLGLLSLAVTSIPSVSNALNWRREFSIQSTLGVAALLISTFHVLIYGWKRAFE 420
QY      421 EBYRRFYTPPNFVLALVLPISIVILDLLQLCRYPD 454
         |||
Db      421 EBYRRFYTPPNFVLALVLPISIVILDLLQLCRYPD 454
```

RESULT 8

```

US-10-455-822-25
; Sequence 25, Application US/10455822
; Publication No. US20040048798A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya
; APPLICANT: Chalilita-Eid, Pia M.
; APPLICANT: Faris, Mary
; TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer
; FILE REFERENCE: 51158-20016.24
; CURRENT APPLICATION NUMBER: US/10/455,822
; PRIOR FILING DATE: 2003-06-04
; PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US60/370,387
; PRIOR FILING DATE: 2002-04-05
```

```

; PRIOR APPLICATION NUMBER: US60/317,840
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: US60/296,656
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US60/091,183
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US60/087,520
; PRIOR FILING DATE: 1998-06-01
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-455-822-25
```

```

Query Match      100.0%; Score 2351; DB 12; Length 454;
Best Local Similarity 100.0%; Pred. No. 6.1e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      1 MESISMGSPKSLSETCLPENGINGIKDARKVTGVTGSGDFAKSLTIRLRGCVHYVIGS 60
         |||
Db      1 MESISMGSPKSLSETCLPENGINGIKDARKVTGVTGSGDFAKSLTIRLRGCVHYVIGS 60
QY      61 RNPKEASEFFPHVVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
         |||
Db      61 RNPKEASEFFPHVVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
QY      121 RINQYPSNAEYLASLFPDSLIVKGFNVSAWALQGPKDASROYVICSNNIQARQVIE 180
         |||
Db      121 RINQYPSNAEYLASLFPDSLIVKGFNVSAWALQGPKDASROYVICSNNIQARQVIE 180
QY      121 RINQYPSNAEYLASLFPDSLIVKGFNVSAWALQGPKDASROYVICSNNIQARQVIE 180
         |||
Db      121 RINQYPSNAEYLASLFPDSLIVKGFNVSAWALQGPKDASROYVICSNNIQARQVIE 180
QY      181 LARQLNFIPIIDLGLSSAREIENLPLRLFTLWRGPVVVAISLATFFFLYSFVRDVIHPYA 240
         |||
Db      181 LARQLNFIPIIDLGLSSAREIENLPLRLFTLWRGPVVVAISLATFFFLYSFVRDVIHPYA 240
QY      181 LARQLNFIPIIDLGLSSAREIENLPLRLFTLWRGPVVVAISLATFFFLYSFVRDVIHPYA 240
         |||
Db      181 LARQLNFIPIIDLGLSSAREIENLPLRLFTLWRGPVVVAISLATFFFLYSFVRDVIHPYA 240
QY      241 RNOQSDFYKPIPIEIVNKTLPPIVATLLSLVYLAGLLAAAYQLYYGTYRRFPFWLETWLO 300
         |||
Db      241 RNOQSDFYKPIPIEIVNKTLPPIVATLLSLVYLAGLLAAAYQLYYGTYRRFPFWLETWLO 300
QY      241 RNOQSDFYKPIPIEIVNKTLPPIVATLLSLVYLAGLLAAAYQLYYGTYRRFPFWLETWLO 300
         |||
Db      241 RNOQSDFYKPIPIEIVNKTLPPIVATLLSLVYLAGLLAAAYQLYYGTYRRFPFWLETWLO 300
QY      301 CRKQLGLISFFPAMVHVAYSLCLPMRSEERYFLNMAYQOVHANINENSWNEEVRRIEMY 360
         |||
Db      301 CRKQLGLISFFPAMVHVAYSLCLPMRSEERYFLNMAYQOVHANINENSWNEEVRRIEMY 360
QY      301 CRKQLGLISFFPAMVHVAYSLCLPMRSEERYFLNMAYQOVHANINENSWNEEVRRIEMY 360
         |||
Db      301 CRKQLGLISFFPAMVHVAYSLCLPMRSEERYFLNMAYQOVHANINENSWNEEVRRIEMY 360
QY      361 ISFGIMSLGLLSLAVTSIPSVSNALNWRREFSIQSTLGVAALLISTFHVLIYGWKRAFE 420
         |||
Db      361 ISFGIMSLGLLSLAVTSIPSVSNALNWRREFSIQSTLGVAALLISTFHVLIYGWKRAFE 420
QY      361 ISFGIMSLGLLSLAVTSIPSVSNALNWRREFSIQSTLGVAALLISTFHVLIYGWKRAFE 420
         |||
Db      361 ISFGIMSLGLLSLAVTSIPSVSNALNWRREFSIQSTLGVAALLISTFHVLIYGWKRAFE 420
QY      421 EBYRRFYTPPNFVLALVLPISIVILDLLQLCRYPD 454
         |||
Db      421 EBYRRFYTPPNFVLALVLPISIVILDLLQLCRYPD 454
```

RESULT 9

```

US-10-455-822-31
; Sequence 31, Application US/10455822
; Publication No. US20040048798A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya
; APPLICANT: Chalilita-Eid, Pia M.
; APPLICANT: Faris, Mary
; TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer
; FILE REFERENCE: 51158-20016.24
; CURRENT APPLICATION NUMBER: US/10/455,822
; PRIOR FILING DATE: 2003-06-04
; PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US60/370,387
; PRIOR FILING DATE: 2002-04-05
```


PRIOR APPLICATION NUMBER: US60/317,840
PRIOR FILING DATE: 2001-09-06
PRIOR APPLICATION NUMBER: US60/296,656
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US60/091,183
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US60/087,520
PRIOR FILING DATE: 1998-06-01
NUMBER OF SEQ ID NOS: 197
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 31
LENGTH: 454
TYPE: PRT
ORGANISM: Homo sapien
US-10-455-822-31

Query Match 100.0%; Score 2351; DB 12; Length 454;
Best Local Similarity 100.0%; Pred. No. 6.1e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVIGS 60
1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVIGS 60
Db 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVIGS 60
QY 61 RNPKFASEFPFHVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLLVGKILIDVSNM 120
61 RNPKFASEFPFHVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLLVGKILIDVSNM 120
Db 61 RNPKFASEFPFHVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLLVGKILIDVSNM 120
QY 121 RINQYPSNAEYLASLFPDGLIVGFNVSAWALQGPKDASROYVICSNNIQARQVIE 180
121 RINQYPSNAEYLASLFPDGLIVGFNVSAWALQGPKDASROYVICSNNIQARQVIE 180
Db 121 RINQYPSNAEYLASLFPDGLIVGFNVSAWALQGPKDASROYVICSNNIQARQVIE 180
QY 181 LARQNFIPIDLGSLSSAREIENLPRLFTLMRGPVVAISLATFFFLYSFYRDVIHPYA 240
181 LARQNFIPIDLGSLSSAREIENLPRLFTLMRGPVVAISLATFFFLYSFYRDVIHPYA 240
Db 181 LARQNFIPIDLGSLSSAREIENLPRLFTLMRGPVVAISLATFFFLYSFYRDVIHPYA 240
QY 241 RNOQSDFYKPIEIVNKTLPVAITLLSLVYLAGLLAAAYQLYYGTYKRRFPWLETWLQ 300
241 RNOQSDFYKPIEIVNKTLPVAITLLSLVYLAGLLAAAYQLYYGTYKRRFPWLETWLQ 300
Db 241 RNOQSDFYKPIEIVNKTLPVAITLLSLVYLAGLLAAAYQLYYGTYKRRFPWLETWLQ 300
QY 301 CRKQGLISFFPAMVHVAISLCLEMRSEERYFLNMAYQOVHANIEENSWNEEYRIEMY 360
301 CRKQGLISFFPAMVHVAISLCLEMRSEERYFLNMAYQOVHANIEENSWNEEYRIEMY 360
Db 301 CRKQGLISFFPAMVHVAISLCLEMRSEERYFLNMAYQOVHANIEENSWNEEYRIEMY 360
QY 361 ISFGIMSLGILLSLAVTSIPSVSNALNMRSEFSFIQSTIGYVALLISTFHVLIYGMKRAFE 420
361 ISFGIMSLGILLSLAVTSIPSVSNALNMRSEFSFIQSTIGYVALLISTFHVLIYGMKRAFE 420
Db 361 ISFGIMSLGILLSLAVTSIPSVSNALNMRSEFSFIQSTIGYVALLISTFHVLIYGMKRAFE 420
QY 421 EEEYRYTPPNFVLALVLPSTIVIDLQLCRYPD 454
421 EEEYRYTPPNFVLALVLPSTIVIDLQLCRYPD 454
Db 421 EEEYRYTPPNFVLALVLPSTIVIDLQLCRYPD 454

RESULT 10
US-10-455-822-33

Sequence 33, Application US/10455822
Publication No. US20040048798A1

GENERAL INFORMATION:

APPLICANT: Agensys, Inc.

APPLICANT: Raitano, Arthur B.

APPLICANT: Ge, Wangmao

APPLICANT: Jakobovits, Aya

APPLICANT: Chalilte-Bid, Pia M.

APPLICANT: Farris, Mary

TITLE OF INVENTION: Nucleic acids and Corresponding Proteins

TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer

FILE REFERENCE: 51158-20016.24

CURRENT APPLICATION NUMBER: US/10/455,822

CURRENT FILING DATE: 2003-06-04

PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned

PRIOR FILING DATE: 2002-12-20

PRIOR APPLICATION NUMBER: US60/370,387

PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
PRIOR FILING DATE: 2001-09-06
PRIOR APPLICATION NUMBER: US60/296,656
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US60/091,183
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US60/087,520
PRIOR FILING DATE: 1998-06-01
NUMBER OF SEQ ID NOS: 197
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 33
LENGTH: 454
TYPE: PRT
ORGANISM: Homo sapien
US-10-455-822-33

Query Match 100.0%; Score 2351; DB 12; Length 454;
Best Local Similarity 100.0%; Pred. No. 6.1e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVIGS 60
1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVIGS 60
Db 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVIGS 60
QY 61 RNPKFASEFPFHVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLLVGKILIDVSNM 120
61 RNPKFASEFPFHVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLLVGKILIDVSNM 120
Db 61 RNPKFASEFPFHVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLLVGKILIDVSNM 120
QY 121 RINQYPSNAEYLASLFPDGLIVGFNVSAWALQGPKDASROYVICSNNIQARQVIE 180
121 RINQYPSNAEYLASLFPDGLIVGFNVSAWALQGPKDASROYVICSNNIQARQVIE 180
Db 121 RINQYPSNAEYLASLFPDGLIVGFNVSAWALQGPKDASROYVICSNNIQARQVIE 180
QY 181 LARQNFIPIDLGSLSSAREIENLPRLFTLMRGPVVAISLATFFFLYSFYRDVIHPYA 240
181 LARQNFIPIDLGSLSSAREIENLPRLFTLMRGPVVAISLATFFFLYSFYRDVIHPYA 240
Db 181 LARQNFIPIDLGSLSSAREIENLPRLFTLMRGPVVAISLATFFFLYSFYRDVIHPYA 240
QY 241 RNOQSDFYKPIEIVNKTLPVAITLLSLVYLAGLLAAAYQLYYGTYKRRFPWLETWLQ 300
241 RNOQSDFYKPIEIVNKTLPVAITLLSLVYLAGLLAAAYQLYYGTYKRRFPWLETWLQ 300
Db 241 RNOQSDFYKPIEIVNKTLPVAITLLSLVYLAGLLAAAYQLYYGTYKRRFPWLETWLQ 300
QY 301 CRKQGLISFFPAMVHVAISLCLEMRSEERYFLNMAYQOVHANIEENSWNEEYRIEMY 360
301 CRKQGLISFFPAMVHVAISLCLEMRSEERYFLNMAYQOVHANIEENSWNEEYRIEMY 360
Db 301 CRKQGLISFFPAMVHVAISLCLEMRSEERYFLNMAYQOVHANIEENSWNEEYRIEMY 360
QY 361 ISFGIMSLGILLSLAVTSIPSVSNALNMRSEFSFIQSTIGYVALLISTFHVLIYGMKRAFE 420
361 ISFGIMSLGILLSLAVTSIPSVSNALNMRSEFSFIQSTIGYVALLISTFHVLIYGMKRAFE 420
Db 361 ISFGIMSLGILLSLAVTSIPSVSNALNMRSEFSFIQSTIGYVALLISTFHVLIYGMKRAFE 420
QY 421 EEEYRYTPPNFVLALVLPSTIVIDLQLCRYPD 454
421 EEEYRYTPPNFVLALVLPSTIVIDLQLCRYPD 454
Db 421 EEEYRYTPPNFVLALVLPSTIVIDLQLCRYPD 454

RESULT 11
US-10-455-822-35

Sequence 35, Application US/10455822
Publication No. US20040048798A1

GENERAL INFORMATION:

APPLICANT: Agensys, Inc.

APPLICANT: Raitano, Arthur B.

APPLICANT: Ge, Wangmao

APPLICANT: Jakobovits, Aya

APPLICANT: Chalilte-Bid, Pia M.

APPLICANT: Farris, Mary

TITLE OF INVENTION: Nucleic acids and Corresponding Proteins

TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer

FILE REFERENCE: 51158-20016.24

CURRENT APPLICATION NUMBER: US/10/455,822

CURRENT FILING DATE: 2003-06-04

PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned

PRIOR FILING DATE: 2002-12-20

PRIOR APPLICATION NUMBER: US60/370,387

PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
PRIOR FILING DATE: 2001-09-06
PRIOR APPLICATION NUMBER: US60/296,656
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US60/091,183
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US60/087,520
PRIOR FILING DATE: 1998-06-01
NUMBER OF SEQ ID NOS: 197
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 35
LENGTH: 454
TYPE: PRT
ORGANISM: Homo sapien
US-10-455-822-35

Query Match 100.0%; Score 2351; DB 12; Length 454;
Best Local Similarity 100.0%; Pred. No. 6.1e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDPAKSLTIRLCGYHVVIGS 60
DB 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDPAKSLTIRLCGYHVVIGS 60
QY 61 RNPKFASEFPFHVDVTHHEDALTKNIIFVAIHREHYTSLMDLRHLLVGKILIDVSNM 120
DB 61 RNPKFASEFPFHVDVTHHEDALTKNIIFVAIHREHYTSLMDLRHLLVGKILIDVSNM 120
QY 121 RINQYPESNAEYLASLPDLSLVKGFNVASAWALQGPKDASROYICSNNIQARQVIE 180
DB 121 RINQYPESNAEYLASLPDLSLVKGFNVASAWALQGPKDASROYICSNNIQARQVIE 180
QY 181 LARQLNFIPIIDGLSLSSAREIENLPLRLFTLWRGPVVVAISLATFFFLYSFVRDVIHPYA 240
DB 181 LARQLNFIPIIDGLSLSSAREIENLPLRLFTLWRGPVVVAISLATFFFLYSFVRDVIHPYA 240
QY 241 RNQOSDFYKIPLEIVNKTLPIVAITLISLVYLAGLLAAAYQLYGTGYRRFPFWLETWLQ 300
DB 241 RNQOSDFYKIPLEIVNKTLPIVAITLISLVYLAGLLAAAYQLYGTGYRRFPFWLETWLQ 300
QY 301 CRKQLGLSFFFAVHVAYSCLPMRSEERYLFNMAYQQVHANINENSWNEEVRRIEMY 360
DB 301 CRKQLGLSFFFAVHVAYSCLPMRSEERYLFNMAYQQVHANINENSWNEEVRRIEMY 360
QY 361 ISFGIMSLGLSLAVTSIPSVSNALNMBEFSFIQSTLGYVALLISTFHVLIYWKRAFE 420
DB 361 ISFGIMSLGLSLAVTSIPSVSNALNMBEFSFIQSTLGYVALLISTFHVLIYWKRAFE 420
QY 421 EBYRFTYTPNFVLAIVLPSIVILDLQLCRYPD 454
DB 421 EBYRFTYTPNFVLAIVLPSIVILDLQLCRYPD 454

RESULT 12
US-10-455-822-37
Sequence 37, Application US/10455822
Publication No. US20040048798A1
GENERAL INFORMATION:
APPLICANT: Agensys, Inc.
APPLICANT: Raitano, Arthur B.
APPLICANT: Ge, Wangmao
APPLICANT: Jakobovits, Aya
APPLICANT: Challita-Eid, Pia M.
APPLICANT: Paris, Mary
TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer
FILE REFERENCE: 51158-20016.24
CURRENT APPLICATION NUMBER: US/10/455,822
CURRENT FILING DATE: 2003-06-04
PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
PRIOR FILING DATE: 2002-12-20
PRIOR APPLICATION NUMBER: US60/370,387
PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
PRIOR FILING DATE: 2001-09-06
PRIOR APPLICATION NUMBER: US60/296,656
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US60/091,183
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US60/087,520
PRIOR FILING DATE: 1998-06-01
NUMBER OF SEQ ID NOS: 197
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 37
LENGTH: 454
TYPE: PRT
ORGANISM: Homo sapien
US-10-455-822-37

Query Match 100.0%; Score 2351; DB 12; Length 454;
Best Local Similarity 100.0%; Pred. No. 6.1e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDPAKSLTIRLCGYHVVIGS 60
DB 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDPAKSLTIRLCGYHVVIGS 60
QY 61 RNPKFASEFPFHVDVTHHEDALTKNIIFVAIHREHYTSLMDLRHLLVGKILIDVSNM 120
DB 61 RNPKFASEFPFHVDVTHHEDALTKNIIFVAIHREHYTSLMDLRHLLVGKILIDVSNM 120
QY 121 RINQYPESNAEYLASLPDLSLVKGFNVASAWALQGPKDASROYICSNNIQARQVIE 180
DB 121 RINQYPESNAEYLASLPDLSLVKGFNVASAWALQGPKDASROYICSNNIQARQVIE 180
QY 181 LARQLNFIPIIDGLSLSSAREIENLPLRLFTLWRGPVVVAISLATFFFLYSFVRDVIHPYA 240
DB 181 LARQLNFIPIIDGLSLSSAREIENLPLRLFTLWRGPVVVAISLATFFFLYSFVRDVIHPYA 240
QY 241 RNQOSDFYKIPLEIVNKTLPIVAITLISLVYLAGLLAAAYQLYGTGYRRFPFWLETWLQ 300
DB 241 RNQOSDFYKIPLEIVNKTLPIVAITLISLVYLAGLLAAAYQLYGTGYRRFPFWLETWLQ 300
QY 301 CRKQLGLSFFFAVHVAYSCLPMRSEERYLFNMAYQQVHANINENSWNEEVRRIEMY 360
DB 301 CRKQLGLSFFFAVHVAYSCLPMRSEERYLFNMAYQQVHANINENSWNEEVRRIEMY 360
QY 361 ISFGIMSLGLSLAVTSIPSVSNALNMBEFSFIQSTLGYVALLISTFHVLIYWKRAFE 420
DB 361 ISFGIMSLGLSLAVTSIPSVSNALNMBEFSFIQSTLGYVALLISTFHVLIYWKRAFE 420
QY 421 EBYRFTYTPNFVLAIVLPSIVILDLQLCRYPD 454
DB 421 EBYRFTYTPNFVLAIVLPSIVILDLQLCRYPD 454

RESULT 13
US-10-455-822-39
Sequence 39, Application US/10455822
Publication No. US20040048798A1
GENERAL INFORMATION:
APPLICANT: Agensys, Inc.
APPLICANT: Raitano, Arthur B.
APPLICANT: Ge, Wangmao
APPLICANT: Jakobovits, Aya
APPLICANT: Challita-Eid, Pia M.
APPLICANT: Paris, Mary
TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer
FILE REFERENCE: 51158-20016.24
CURRENT APPLICATION NUMBER: US/10/455,822
CURRENT FILING DATE: 2003-06-04
PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
PRIOR FILING DATE: 2002-12-20
PRIOR APPLICATION NUMBER: US60/370,387
PRIOR FILING DATE: 2002-04-05


```

; PRIOR APPLICATION NUMBER: US60/317,840
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: US60/296,656
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US60/091,183
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US60/087,520
; PRIOR FILING DATE: 1998-06-01
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 39
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-455-822-39

```

```

Query Match      100.0%; Score 2351; DB 12; Length 454;
Best Local Similarity 100.0%; Pred. No. 6.1e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 MESISMGSPKSLSETCPLNGINGIKDARKVTGVIGSGDPAKSLTIRLCGYHVIGS 60
DB      1 MESISMGSPKSLSETCPLNGINGIKDARKVTGVIGSGDPAKSLTIRLCGYHVIGS 60
QY      61 RNPKFASEFFPHVVDVTHHEDALTNTNIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
DB      61 RNPKFASEFFPHVVDVTHHEDALTNTNIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
QY      121 RINQYPESNAEYLASLPDLSLVKGFNVVSAMALQGPKDASROVYICSNNIQARQVIE 180
DB      121 RINQYPESNAEYLASLPDLSLVKGFNVVSAMALQGPKDASROVYICSNNIQARQVIE 180
QY      181 LARQNFIPIDLSLSAREIENLPRLFTLWGPVVVAISLATFFFLYSFVRDVIHPYA 240
DB      181 LARQNFIPIDLSLSAREIENLPRLFTLWGPVVVAISLATFFFLYSFVRDVIHPYA 240
QY      241 RNOQSDFYKIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYGTXYRRPFWLETWLQ 300
DB      241 RNOQSDFYKIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYGTXYRRPFWLETWLQ 300
QY      301 CRKQGLLSFFFAMVHVAVSLCLPMRRESRYLFLNMAVYQOVHANINENSNNEEVRRIEMY 360
DB      301 CRKQGLLSFFFAMVHVAVSLCLPMRRESRYLFLNMAVYQOVHANINENSNNEEVRRIEMY 360
QY      361 ISFGIMSLGLSLLAVTSIPSVSNALNWRREFSFIOSTLGYVALLISTFHVLIYGWKRAFE 420
DB      361 ISFGIMSLGLSLLAVTSIPSVSNALNWRREFSFIOSTLGYVALLISTFHVLIYGWKRAFE 420
QY      421 EEYRRFYTPPNFVALVLPISIVILDLLQLCRYPD 454
DB      421 EEYRRFYTPPNFVALVLPISIVILDLLQLCRYPD 454

```

```

RESULT 14
US-10-455-822-78
; Sequence 78, Application US/10455822
; Publication No. US20040048798A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Farris, Mary
; TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 98PA86 Useful in Treatment and Detection of Cancer
; FILE REFERENCE: 51158-20016.24
; CURRENT APPLICATION NUMBER: US/10/455,822
; CURRENT FILING DATE: 2003-06-04
; PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US60/370,387
; PRIOR FILING DATE: 2002-04-05

```

```

; PRIOR APPLICATION NUMBER: US60/317,840
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: US60/296,656
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US60/091,183
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US60/087,520
; PRIOR FILING DATE: 1998-06-01
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 78
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-455-822-78

```

```

Query Match      100.0%; Score 2351; DB 12; Length 454;
Best Local Similarity 100.0%; Pred. No. 6.1e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 MESISMGSPKSLSETCPLNGINGIKDARKVTGVIGSGDPAKSLTIRLCGYHVIGS 60
DB      1 MESISMGSPKSLSETCPLNGINGIKDARKVTGVIGSGDPAKSLTIRLCGYHVIGS 60
QY      61 RNPKFASEFFPHVVDVTHHEDALTNTNIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
DB      61 RNPKFASEFFPHVVDVTHHEDALTNTNIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
QY      121 RINQYPESNAEYLASLPDLSLVKGFNVVSAMALQGPKDASROVYICSNNIQARQVIE 180
DB      121 RINQYPESNAEYLASLPDLSLVKGFNVVSAMALQGPKDASROVYICSNNIQARQVIE 180
QY      181 LARQNFIPIDLSLSAREIENLPRLFTLWGPVVVAISLATFFFLYSFVRDVIHPYA 240
DB      181 LARQNFIPIDLSLSAREIENLPRLFTLWGPVVVAISLATFFFLYSFVRDVIHPYA 240
QY      241 RNOQSDFYKIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYGTXYRRPFWLETWLQ 300
DB      241 RNOQSDFYKIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYGTXYRRPFWLETWLQ 300
QY      301 CRKQGLLSFFFAMVHVAVSLCLPMRRESRYLFLNMAVYQOVHANINENSNNEEVRRIEMY 360
DB      301 CRKQGLLSFFFAMVHVAVSLCLPMRRESRYLFLNMAVYQOVHANINENSNNEEVRRIEMY 360
QY      361 ISFGIMSLGLSLLAVTSIPSVSNALNWRREFSFIOSTLGYVALLISTFHVLIYGWKRAFE 420
DB      361 ISFGIMSLGLSLLAVTSIPSVSNALNWRREFSFIOSTLGYVALLISTFHVLIYGWKRAFE 420
QY      421 EEYRRFYTPPNFVALVLPISIVILDLLQLCRYPD 454
DB      421 EEYRRFYTPPNFVALVLPISIVILDLLQLCRYPD 454

```

```

RESULT 15
US-10-455-822-121
; Sequence 121, Application US/10455822
; Publication No. US20040048798A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovits, Aya
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Farris, Mary
; TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
; TITLE OF INVENTION: Entitled 98PA86 Useful in Treatment and Detection of Cancer
; FILE REFERENCE: 51158-20016.24
; CURRENT APPLICATION NUMBER: US/10/455,822
; CURRENT FILING DATE: 2003-06-04
; PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US60/370,387
; PRIOR FILING DATE: 2002-04-05

```

; PRIOR APPLICATION NUMBER: US60/317,840
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: US60/296,656
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US60/091,183
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US60/087,520
; PRIOR FILING DATE: 1998-06-01
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 121
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-455-822-121

Query Match 100.0%; Score 2351; DB 12; Length 454;
Best Local Similarity 100.0%; Pred. No. 6.1e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGMSPKSLSETCLPNGINGIKDARKYTVGIVGSGDFAKSLTIRLRGCVHVVIGS 60
Db 1 MESISMGMSPKSLSETCLPNGINGIKDARKYTVGIVGSGDFAKSLTIRLRGCVHVVIGS 60
QY 61 RNPKFASEFFPHVVDVTHHEDALTKNIIFVAIHREHYTSLWDLRHLLVGKILLIDVSNM 120
Db 61 RNPKFASEFFPHVVDVTHHEDALTKNIIFVAIHREHYTSLWDLRHLLVGKILLIDVSNM 120
QY 121 RINQYPESNAEYLAGLFPDLSLVKGFNVVSAMALQLGPKDASRQVYICSNNIQARQVIE 180
Db 121 RINQYPESNAEYLAGLFPDLSLVKGFNVVSAMALQLGPKDASRQVYICSNNIQARQVIE 180
QY 181 LARQLNFIPIDLGSLSSAREIENLPLRLFTLMRGVVAISLATFFFLYSFVRDVIHPYA 240
Db 181 LARQLNFIPIDLGSLSSAREIENLPLRLFTLMRGVVAISLATFFFLYSFVRDVIHPYA 240
QY 241 RNQOSDFYKIPIEIVNKTLPVIAITLLSLVYLAGLLAAAYQLYGTKYRRFPWLETWLQ 300
Db 241 RNQOSDFYKIPIEIVNKTLPVIAITLLSLVYLAGLLAAAYQLYGTKYRRFPWLETWLQ 300
QY 301 CRKQLGLSFFFAMVHVAYSLCLPMRSEKYLPNMAVQOVHANIENTSNWNEEVRIMY 360
Db 301 CRKQLGLSFFFAMVHVAYSLCLPMRSEKYLPNMAVQOVHANIENTSNWNEEVRIMY 360
QY 361 ISFGIMSLGLSLLAVTSLPSVSNALNWRREFSFIOSTLGYVALLISTFHVLIYGWKRAFE 420
Db 361 ISFGIMSLGLSLLAVTSLPSVSNALNWRREFSFIOSTLGYVALLISTFHVLIYGWKRAFE 420
QY 421 EBYRFRYTPPNFVALVLPISIVLIDLQLCRYPD 454
Db 421 EBYRFRYTPPNFVALVLPISIVLIDLQLCRYPD 454

Search completed: April 16, 2004, 12:46:36
Job time : 49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 12:38:02 ; Search time 22 Seconds
(without alignments)
1065.372 Million cell updates/sec

Title: US-09-455-486-6
Perfect score: 2351
Sequence: 1 MESISMGMGSPKSLSETCLPN.....ALVLPISIVLDLQLCRYPD 454

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfill1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	901	38.3	173	4	US-09-323-873A-8
2	736	31.3	141	3	US-09-083-521-1
3	717	30.5	339	4	US-09-323-873A-2
4	717	30.5	339	4	US-09-685-166A-879
5	144.5	6.1	227	4	US-09-655-270A-15
6	144.5	6.1	227	4	US-09-651-941-17
7	144.5	6.1	227	4	US-09-955-597-17
8	107.5	4.6	695	1	US-08-487-886-2
9	107.5	4.6	695	3	US-08-482-855-2
10	107.5	4.6	695	4	US-08-474-986-2
11	107	4.6	34	4	US-09-323-873A-20
12	102.5	4.4	940	4	US-09-328-352-8165
13	101	4.3	476	3	US-09-316-083-3
14	101	4.3	476	4	US-09-933-700-3
15	101	4.3	724	4	US-09-252-991A-21494
16	99	4.2	365	4	US-09-170-496D-118
17	99	4.2	365	4	US-09-170-496D-226
18	99	4.2	365	4	US-09-364-425B-27
19	99	4.2	531	2	US-08-724-974A-2
20	97.5	4.1	390	3	US-08-460-576-2
21	97	4.1	692	3	US-07-757-342D-6
22	97	4.1	692	4	US-09-461-657B-6
23	96	4.1	299	4	US-09-903-456-61
24	96	4.1	1309	4	US-09-975-413A-10
25	94.5	4.0	288	4	US-09-107-532A-6009
26	94.5	4.0	296	4	US-09-540-236-2922
27	94.5	4.0	345	4	US-09-489-039A-10740

28	94	4.0	284	4	US-09-903-456-83	Sequence 83, Appl
29	94	4.0	292	4	US-09-903-456-35	Sequence 35, Appl
30	94	4.0	293	4	US-09-903-456-45	Sequence 45, Appl
31	93	4.0	365	2	US-08-724-974A-3	Sequence 3, Appl
32	93	4.0	498	4	US-09-107-532A-7077	Sequence 7077, Ap
33	92.5	3.9	405	4	US-09-489-039A-9411	Sequence 9411, Ap
34	91.5	3.9	407	4	US-09-328-352-5605	Sequence 5605, Ap
35	91.5	3.9	495	4	US-08-985-343-2	Sequence 2, Appl
36	91	3.9	211	4	US-09-252-991A-23822	Sequence 23822, A
37	91	3.9	440	4	US-09-634-238-281	Sequence 281, App
38	91	3.9	550	4	US-09-543-681A-6026	Sequence 6026, App
39	91	3.9	808	4	US-09-134-001C-3105	Sequence 3105, Ap
40	90.5	3.8	340	4	US-09-543-681A-7503	Sequence 7503, Ap
41	90.5	3.8	388	4	US-09-107-532A-6381	Sequence 6381, Ap
42	90	3.8	500	4	US-09-178-093B-26	Sequence 26, Appl
43	90	3.8	970	4	US-09-795-927-7	Sequence 7, Appl
44	89.5	3.8	495	1	US-07-841-997A-2	Sequence 2, Appl
45	89.5	3.8	495	1	US-08-290-301-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-323-873A-8
; Sequence 8, Application US/09323873A
; Patent No. 6329503
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Kahar Leong
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Steve Chappell Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
; FILE REFERENCE: 129.16USU2
; CURRENT APPLICATION NUMBER: US/09/323,873A
; CURRENT FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-09-323-873A-8

Query Match 38.3%; Score 901; DB 4; Length 173;
Best Local Similarity 100.0%; Pred. No. 1.4e-86;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 246 DFYKIPFIEIVNKTPIVAITLTLSTLVYLAGLLAAAYOLYGTGKRRFPFWLETWLOCRKOL 305
DB 1 DFYKIPFIEIVNKTPIVAITLTLSTLVYLAGLLAAAYOLYGTGKRRFPFWLETWLOCRKOL 60

QY 306 GLISFFPAMVHVAYSICLPMRSEKYLFLNNAYQOVHANIEINSWNEEYWRIEYISFGI 365
DB 61 GLISFFPAMVHVAYSICLPMRSEKYLFLNNAYQOVHANIEINSWNEEYWRIEYISFGI 120

QY 366 MSIGLLSLAVTSIPSVSNALNWRFFSFIQSTLGYVALLISTFHVLIYWKRA 418
DB 121 MSIGLLSLAVTSIPSVSNALNWRFFSFIQSTLGYVALLISTFHVLIYWKRA 173

RESULT 2
US-09-083-521-1
; Sequence 1, Application US/09083521
; Patent No. 6048970
; GENERAL INFORMATION:

Fri Apr 16 16:26:35 2004

us-09-455-486-6.raj

Page 2

APPLICANT: Lal, Preeti
 APPLICANT: Guegler, Karl J.
 APPLICANT: Corley, Neil C.
 TITLE OF INVENTION: PROSTATE GROWTH-ASSOCIATED MEMBRANE PROTEINS
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 STREET: 3174 PORTER DRIVE
 CITY: PALO ALTO
 STATE: CALIFORNIA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: word Perfect 6.1 for Windows/MS-DOS 6.2
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/083,521
 FILING DATE: Herewith
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: CERRONE, MICHAEL C.
 REGISTRATION NUMBER: 39,132
 REFERENCE/DOCKET NUMBER: PF-0527 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (650) 855-0555
 TELEFAX: (650) 845-4166
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 141 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: PROSTUT10
 CLONE: 1691243

```

Query Match          31.3%; Score 736; DB 3; Length 141;
Best Local Similarity 100.0%; Pred. No. 2.2e-69;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      314  MVHVAYSLCLPMRRSERYLFLNNAYQQVHANIEINSWNEEEVWRIEMYSISFGIMSLGLSL 373
      |||
      1  MHVAYSLCLPMRRSERYLFNNAYQQVHANIEINSWNEEEVWRIEMYSISFGIMSLGLSL 60

QY      374  LAVTSPSVSNALNMRREFSFIQSTLGYVALLISTFHVLIYGWKRAFEEEYRFPYTPPNFV 433
      |||
      61  LAVTSPSVSNALNMRREFSFIQSTLGYVALLISTFHVLIYGWKRAFEEEYRFPYTPPNFV 120

DB      434  LALVLPISIVILDLQLCRYPD 454
      |||
      121  LALVLPISIVILDLQLCRYPD 141

RESULT 3
US-09-323-873A-2
; Sequence 2, Application US/09323873A
; Patent No. 6329503
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Kahan Leong
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Steve Chappell Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
; FILE REFERENCE: 129.16USU2
; CURRENT APPLICATION NUMBER: US/09/323, 873A
; CURRENT FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520

```

```

; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo Sapiens
;
US-09-323-873A-2

```

Query Match	30.5%;	Score 717;	DB 4;	Length 339;
Best Local Similarity	54.9%;	Pred. No. 8.5e-67;		
Matches 130; Conservative	48;	Mismatches 59;	Indels 0;	Gaps 0;

[illegible]

RESULT 4
US-09-685-166A-879
Sequence 879, Application US/09685166A

```

; Patent NO: 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqun
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-685-166A-879

```

	Query Match	30.5%;	Score 717;	DB 4;	Length 339;	
	Best Local Similarity	54.9%;	Pred. No. 8.5e-67;			
	Matches 130;	Conservative 48;	Mismatches 59;	Indels 0;	Gaps 0	
QY	208 LFTLMRGDVVAISLATFFFLYSFVRDVIHPYARNQQSDPKPIPIEIVNKTLPVATILL	267				
	: : : : : :					


```

Sequence 2, Application US/08487886
Patent No. 5744448
GENERAL INFORMATION:
APPLICANT: Kelton, Christie Ann
APPLICANT: Schweickhardt, Rene Lynn
APPLICANT: Cheng, Shirley Vui Yen
APPLICANT: Nugent, No. 5744448een Patrice
TITLE OF INVENTION: Human Pollicle Stimulating
TITLE OF INVENTION: Hormone Receptor
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Stephan P. Williams,
ADDRESS: Ares-Serono, Inc.
STREET: Exchange Place, 37th floor
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
COMPUTER: IBM PS/2, model 55 SX
OPERATING SYSTEM: MS-DOS version 4.0
SOFTWARE: VAX/VMS Masell via Kermit to IBM MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,886
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/670,085
FILING DATE: 15-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Stephan P.
REGISTRATION/DOCKET NUMBER: 28546
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 723-1300
TELEFAX: (617) 723-8923
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 695
TYPE: Amino acid
TOPOLOGY: Linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: signal sequence
LOCATION: -17 to -1
IDENTIFICATION METHOD: hydrophobic
FEATURE:
NAME/KEY: putative amino-terminal extracellular domain
LOCATION: 1 to 349
IDENTIFICATION METHOD: similarity with other
IDENTIFICATION METHOD: dimeric glycoprotein receptor extracellular
IDENTIFICATION METHOD: domains, hydrophilic
FEATURE:
NAME/KEY: transmembrane domain
LOCATION: 350 to 613
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane domains
FEATURE:
NAME/KEY: putative transmembrane region I
LOCATION: 350 to 370
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region II
LOCATION: 382 to 404
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region III
LOCATION: 427 to 448

```

```

IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region IV
LOCATION: 469 to 491
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region V
LOCATION: 512 to 533
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region VI
LOCATION: 557 to 580
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region VII
LOCATION: 592 to 613
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative carboxy-terminal intracellular
LOCATION: 614 to 678
US-08-487-886-2

Query Match 4.6%; Score 107.5; DB 1; Length 695;
Best Local Similarity 17.0%; Pred. No. 0.028;
Matches 84; Conservative 77; Mismatches 148; Indels 185; Gaps 17;

QY 79 HEDALTKNTIIFVAIHREHYTSLMDRLHLVG-----KILIDVSNMR 121
DB 98 HEIRIEKANL-LYINPEAFONLPNLQYLLISNTGKHLPDVHKIHSLOKVLDDIQDNIN 156
QY 122 INQYPSNAEYLASLPDGLIVKGFNVSAWALQCPKASROVYICSNINQARQVIEL 181
DB 157 IH-----TIERNSEYGLSESVILWL---NKGIOEIHNC----- 189
QY 182 ARQLNFPIDGLSSAREIENLPLFTLMRGPVVAISLATFFLYSFVRVYHPYAR 241
DB 190 ---FNGQJDELNLSDNNLELPNDVFHGASGPVILDISRTIRHSLPSYGLENLKIRA 246
QY 242 NQSDFYKPIEIVNKTLP-PIVAITLLSLVY----- 271
DB 247 RSTYNLKKLP-----TLKXVAMEASLTTPSHCCAFANMRQISELHPICNKSILROE 300
QY 272 -----LAGLLAAAYQLYYGTKYRRF-----PPWLETWLQCRKOL 305
DB 301 VDYMTQTRGQRSSLAEDNESSYGRGDMTYTEFDYDLCNEVYDVTCSPKPDAFNPEEDIM 360
QY 306 G-----LSFFFAVHVA-----YSLCLPMRBSERYLFLNMAVQOVHANTEN 347
DB 361 GYNILRVLIWFISILAITGNIIVVILTSQYKLTVP-----RFLMCNLAFAD----- 408
QY 348 SWNEEVRRIEMYISFGIMSLGLSLAVTSIPSVSNALNRE-----FSFIQSTL 398
DB 409 -----LCIGYLLILASVDIHTKSQYHNVAIDWGTAGCDAAGFTVFASEL 455
QY 399 G---YVALISTFH-----VLIYGWKRAFESEYRYRFTPPNFV-LA 435
DB 456 SVYTLTAITLERWHTITHAMQLDCKVQLRHAASVVMWGMIFAALFPIFGISSYMKVS 515
QY 436 LVLPSTIVLDLQL 449
DB 516 ICLPMDIDSPLSQL 529

```

```

RESULT 9
US-08-482-855-2
Sequence 2, Application US/08482855
Patent No. 6121016
GENERAL INFORMATION:
APPLICANT: Kelton, Christie Ann
APPLICANT: Schweickhardt, Rene Lynn
APPLICANT: Cheng, Shirley Vui Yen
APPLICANT: Nugent, No. 6121016een Patrice
TITLE OF INVENTION: Human Pollicle Stimulating
TITLE OF INVENTION: Hormone Receptor
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Stephan P. Williams,
ADDRESSEE: Areo-Serono, Inc.
STREET: Exchange Place, 37th floor
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
COMPUTER: IBM PS/2, model 55 SX
OPERATING SYSTEM: MS-DOS version 4.0
SOFTWARE: VAX/VMS Massll via Kermit to IBM MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,855
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/670,085
FILING DATE: 15-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Stephan P.
REGISTRATION NUMBER: 28546
REFERENCE/DOCKET NUMBER: US/252
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 723-1300
TELEFAX: (617) 723-8923
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 695
TYPE: Amino acid
TOPOLOGY: Linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: signal sequence
LOCATION: -17 to -1
IDENTIFICATION METHOD: hydrophobic
FEATURE:
NAME/KEY: putative amino-terminal extracellular domain
LOCATION: 1 to 349
IDENTIFICATION METHOD: similarity with other
IDENTIFICATION METHOD: dimeric glycoprotein receptor extracellular
IDENTIFICATION METHOD: domains, hydrophilic
FEATURE:
NAME/KEY: transmembrane domain
LOCATION: 350 to 613
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane domains
FEATURE:
NAME/KEY: putative transmembrane region I
LOCATION: 350 to 370
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region II
LOCATION: 382 to 404
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length

```

```

FEATURE:
NAME/KEY: putative transmembrane region III
LOCATION: 427 to 448
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region IV
LOCATION: 469 to 491
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region V
LOCATION: 512 to 533
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region VI
LOCATION: 557 to 580
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region VII
LOCATION: 592 to 613
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative carboxy-terminal intracellular
LOCATION: 614 to 678
US-08-482-855-2

Query Match 4.6%; Score 107.5; DB 3; Length 695;
Best Local Similarity 17.0%; Pred. No. 0.028;
Matches 84; Conservative 77; Mismatches 148; Indels 185; Gaps 17;

QY 79 HEDALTKNTIIFVAIHREHYTSLMDLRHLVG-----KILIVSNMR 121
DB 98 HEIRTEKANNL-LYINPEAFQNLPLQYLLISNTGIRKLPDVHKIHSLOKVLIDQDNIN 156
QY 122 INQYPESNAEYIASLFPDSLIVKGFNVSAWALQIGPKDASROVYICSNINQARQVIEL 181
DB 157 IH-----TIERSEFVGLSFESVILWL-----NKGIOEIHNCA----- 189
QY 182 ARQLNFIPIIDGLSSAREIENLPLRFTLWRGPVVAISLATFFLYSFVRDVIHPYAR 241
DB 190 ---FNGQLDELNLSDNNMLELPNDVFHGASGPVILDISRTRIHSLPSYGLENIKLIRA 246
QY 242 NQGSDFYKIPLEIVNKTLLPIVAITLLSLVY----- 271
DB 247 RSTYNLKLPL-----TLEKVALMEASLTYPSSHCAFWNRQISELHPICNKSILROE 300
QY 272 -----LAGLLAAVQLYYGTYRRF-----PWLFTWLQCRKOL 305
DB 301 VDYMTQTRGQSSLAEDNESSYRGFDMTYTEFDYDLCEVVDVTCSPKPDAFNPEDIM 360
QY 306 G-----LSFFFAVHVA-----YSICLPMRSEERYFLNMAVQCVHANIE 347
DB 361 GYNILRYLWIFISILAITGNITVLITTSQYKLTVP-----RFLMCNLAFAD----- 408
QY 348 SMNEEWEVRIEMYISFGIMSLGLSLAVTSIPSVSNALNMR-----FSFIQSTL 398
DB 409 -----LCIGIYLLIASVDIHTKSQYHNVAIDMQTGAGCDAAAGFTVFASEL 455
QY 399 G---YVALLISTFH-----VLIYGNKRAFESEYRYFYTPPNFV-LA 435
DB 456 SVYTLTALITERWHTITHAMQLDCKVQLRHAASVVMGWIFAFAALFPIFGISSYMKVS 515
QY 436 LVLPSTIVLDLQL 449

```



```

QY 88 IIFVAIHREHYTSLMDLRHLVKGKILIDVSNMNRINQYPSNAEYLSL-----FPDSL1 142
DB 262 VVFLTLVFSHFGV--ITSLTTASILFTILAIITVFLSLKQQAITYLAIALGMAYAAPLV 319
QY 143 VKGF--NVWSAWALQLGPKDASRQVYICSNNIQARQOVIELARQIN---FIPIDGLSL 196
DB 320 IPQYRPDVVFLFS-----YYLVINLAVAANFIQPKILNQIAFFATMFIGSA 368
QY 197 SAREIENTPLRLFTL-WRGPVVAISLAFEPFLYSFVRDVIHPYARNQOSDFYKIPIEIV 255
DB 369 IAFYAE--PAKEDTLDM---ILMLHIALFTLWLSVRYSQNISRSVSEHEKQEGIRLP---- 418
QY 256 NKTLPVIAITLISLVYLAGLLAAAYQLYGTGK-----YRRFPWLETWLQCRK 303
DB 419 ----PLDVGLEFNVFVLGFTLHAYLVHESQALTTGAAVLAGTYAVLTFWIK--KTHP 471
QY 304 QLGLLSFFFAWVHA-YSLCLPMRSERYLFLMAYQQOVHANIENSNBEEVWRIEMYIS 362
DB 472 QLSVLAKSFFILAVAFALIIFPLAKGAHWTAIGWAQGTALIV--WGVTBRYSRLSRYI- 527
QY 363 FGIMSLGLSLAVTISPSVSNALNW--REFSFIQSTLGVALLSTFHVLIYGMKRAFE 420
DB 528 -GVILVLLSSLALFYQV-----WANEFFPLTSTSIYIAQFISAFYLLQYNSK---E 575
QY 421 EBY 424
DB 576 QRYF 579

```

```

RESULT 13
US-09-316-083-3
; Sequence 3, Application US/09316083A
; Patent No. 6280942
; GENERAL INFORMATION:
; APPLICANT: The Institute of Physical and Chemical Research
; TITLE OF INVENTION: Endonuclease
; FILE REFERENCE: PH-651
; CURRENT APPLICATION NUMBER: US/09/316, 083A
; CURRENT FILING DATE: 1999-05-20
; EARLIER APPLICATION NUMBER: JP98/141861
; EARLIER FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-316-083-3

```

```

Query Match 4.3%; Score 101; DB 3; Length 476;
Best Local Similarity 20.7%; Pred. No. 0.075;
Matches 52; Conservative 48; Mismatches 85; Indels 66; Gaps 12;

QY 62 NPKFASFPFPHVVDVTHHEDALTKTNI---IFVAIHREHYT-----SLWD 103
DB 254 NPYFVNAFSINI-----KTNLAKEKIFPTNIYNKLYSDYKINQINNHIPYNYLK 302
QY 104 LRHLVKGKILIDVSNMNRINQYPSNAEYLSL--PDSLIVKGFNVSAWALQLGPKDA 161
DB 303 INNKLPKINIMDKNNYWLAGTAAAGSFLSSMYNPKDITLTKNM----- 347
QY 162 SRQVYICSNNIQARQOVIELARQINFIPIDLGSLSSAREIENTPLRLFTLWRGPVVAIS 221
DB 348 -RPSYVIS-QVETRKEILYIQE---SFDL-SISNVKVGNRKLDKFKLFTRTTDELMK 400
QY 222 LATFEPFLYSFVRDVIHPYARNQOSDFYKIP---IEIVNKTLPVIAITLISLVYLAGLLA 277
DB 401 ----FIYYF--DKFLPLHDNKQFNVIKFRFNTFIKSYWNNRVFGLVLSB--YINNIXI 451
QY 278 AAYQLYGTGY 288
DB 452 DNYDYYYNKKY 462

```

```

RESULT 14
US-09-933-700-3
; Sequence 3, Application US/099333700
; Patent No. 6528296
; GENERAL INFORMATION:
; APPLICANT: The Institute of Physical and Chemical Research
; TITLE OF INVENTION: Endonuclease
; FILE REFERENCE: PH-651
; CURRENT APPLICATION NUMBER: US/09/933,700
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 09/316,083
; PRIOR FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-933-700-3

```

```

Query Match 4.3%; Score 101; DB 4; Length 476;
Best Local Similarity 20.7%; Pred. No. 0.075;
Matches 52; Conservative 48; Mismatches 85; Indels 66; Gaps 12;

QY 62 NPKFASFPFPHVVDVTHHEDALTKTNI---IFVAIHREHYT-----SLWD 103
DB 254 NPYFVNAFSINI-----KTNLAKEKIFPTNIYNKLYSDYKINQINNHIPYNYLK 302
QY 104 LRHLVKGKILIDVSNMNRINQYPSNAEYLSL--PDSLIVKGFNVSAWALQLGPKDA 161
DB 303 INNKLPKINIMDKNNYWLAGTAAAGSFLSSMYNPKDITLTKNM----- 347
QY 162 SRQVYICSNNIQARQOVIELARQINFIPIDLGSLSSAREIENTPLRLFTLWRGPVVAIS 221
DB 348 -RPSYVIS-QVETRKEILYIQE---SFDL-SISNVKVGNRKLDKFKLFTRTTDELMK 400
QY 222 LATFEPFLYSFVRDVIHPYARNQOSDFYKIP---IEIVNKTLPVIAITLISLVYLAGLLA 277
DB 401 ----FIYYF--DKFLPLHDNKQFNVIKFRFNTFIKSYWNNRVFGLVLSB--YINNIXI 451
QY 278 AAYQLYGTGY 288
DB 452 DNYDYYYNKKY 462

```

```

RESULT 15
US-09-252-991A-21494
; Sequence 21494, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21494
; LENGTH: 724
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21494

```

```

Query Match 4.3%; Score 101; DB 4; Length 724;
Best Local Similarity 20.5%; Pred. No. 0.14;
Matches 72; Conservative 57; Mismatches 105; Indels 118; Gaps 16;

```

QY 132 YLASLFPDLSLVKGFNVVSAMALQLGPKDASRQVYICSNNIQA-----RQVIELARQLN 186
Db 298 YLSVVLPPDSL-SKTLAMWAVVLVCGTLFSA----LCVISLSLSGPHRQRALDILRRQA 352
QY 187 FIPIDL-GSLSSAREIENLPLRLFTLWRGPVVAISLA-----TFEFLYSFVRDVIHP 238
Db 353 FRPLWLIGSLALGEVAHDPRLLIAGLGEHTSICLSTLANASAALEFVALFWMRFRRIAH- 411
QY 239 YARNQO-----SDFYKIPi-----EIVNK 257
Db 412 LIRNQLERRLKRRSLHDLVQLVGSIMFVPLVLVGISLFATFVSAGDSSALRRALVCA 471
QY 258 TLPIVAITLLSLVYLAGLLAAAYQLYGTKYRRFPFMLETWLQCRKQGLLSEFFAMVHV 317
Db 472 VLAVVAMTVIGLI-----RRSSRVGAGPRRSAPYIBQ-----LQSFGYTLIHI 515
QY 318 AYSLCLPMRRSERYLFLNMAVYQOVHANIENTSNWNEEVRRIEM-YISFGIM----- 366
Db 516 FVVL-----FFIEVALRWGMSLIR-YAEGEGEQISMKVVSFGTLLVAMLWIL 564
QY 367 -----SLGL-----LSLLAVTSIPSVSNALNWRKESFIQSTLGYVALLIS 406
Db 565 TDTAIOHSLGLGKSRPNTRALTMLPLIRNVL-----FATIAVIALIVA 608

Search completed: April 16, 2004, 12:41:47
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 12:32:36 ; Search time 60 Seconds
(without alignments)
2137.943 Million cell updates/sec

Title: US-09-455-486-6
Perfect score: 2351
Sequence: 1 MESISMNGSPKSLSETCLPN.....ALVLPSTVLDLQLCRYPD 454

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2351	100.0	454	4	AAE02781	Aae02781 Human six
2	2351	100.0	454	4	AAU10188	Aau10188 Human ORF
3	2351	100.0	454	5	ABG61933	Abg61933 Prostate
4	2351	100.0	454	5	AAU76538	Aau76538 Tumour-as
5	2351	100.0	454	5	AAU80190	Aau80190 Human PUM
6	2351	100.0	454	7	ABU08893	Abu08893 Tumour-as
7	2348	99.9	454	7	ADB65001	Adb65001 Human pro
8	2294	97.6	490	5	AAE28951	Aae28951 Human STR
9	2294	97.6	490	7	ABU63312	Abu63312 Human tra
10	2290	97.4	490	4	AAU10187	Aau10187 Human six
11	2036	86.6	419	4	AAU10189	Aau10189 Human ORF
12	1972.5	83.9	576	4	ABG12306	Abg12306 Novel hum
13	1972.5	83.9	1273	4	ABG00113	Abg00113 Novel hum
14	1738	73.9	450	4	AAE02841	Aae02841 Human STR
15	1364.5	58.0	1082	5	ABP62883	Abp62883 Human pol
16	1316	56.0	488	4	AAAB49483	Aab49483 Rat p-HYD
17	1316	56.0	526	5	ABB83365	Abb83365 Murine Tu
18	1272	54.1	488	4	AAAB85775	Aab85775 Human dru
19	1272	54.1	488	5	ABB83366	Abb83366 Human Tum
20	1269	54.0	488	4	AAAB93224	Aab93224 Human pro
21	1269	54.0	488	4	AAU04564	Aau04564 Human G-p
22	1269	54.0	488	4	AAU10220	Aau10220 Human Six
23	1269	54.0	488	6	ABU60864	Abu60864 Human G p
24	1245.5	53.0	487	4	AAAB49481	Aab49481 Human p-H
25	1088	46.3	459	4	AAAB74715	Aab74715 Human mem

26	1087	46.2	459	4	AAE02782	Aae02782 Human six
27	1087	46.2	459	5	ABP64820	Abp64820 Human pro
28	1085	46.2	459	4	AAU10190	Aau10190 Human Six
29	1082	46.0	458	4	AAE02636	Aae02636 Human STR
30	1074.5	45.7	456	4	AAAB49482	Aab49482 Human p-H
31	901	38.3	173	3	AAAY58195	Aay58195 Human STR
32	844	35.9	179	4	AAAM18006	Aam18006 Peptide #
33	844	35.9	179	4	ABBB37041	Abb37041 Peptide #
34	844	35.9	179	4	AAAM30517	Aam30517 Peptide #
35	844	35.9	179	4	ABBB31810	Abb31810 Peptide #
36	844	35.9	179	4	ABB22356	Abb22356 Protein #
37	844	35.9	179	4	AAAM70183	Aam70183 Human bon
38	844	35.9	179	4	AAAM57768	Aam57768 Human bra
39	844	35.9	179	4	ABG51883	Abg51883 Human liv
40	844	35.9	179	4	AAAM05646	Aam05646 Peptide #
41	844	35.9	179	5	ABG39817	Abg39817 Human pep
42	768	32.7	237	4	AAU04565	Aau04565 Human G-p
43	768	32.7	237	6	ABU60865	Abu60865 Human G p
44	736	31.3	141	3	AAAY52589	Aay52589 Human pro
45	717	30.5	267	6	ABU60886	Abu60886 Human G p

ALIGNMENTS

RESULT 1	
AAE02781	
ID	AAE02781 standard; protein; 454 AA.
XX	AC
XX	AAE02781;
DT	06-AUG-2001 (first entry)
XX	DE
XX	Human six transmembrane epithelial antigen of prostate (STEAP)-2 protein.
KW	Human; cytosolic; antiproliferative; vaccine; gene therapy;
KW	six transmembrane epithelial antigen of the prostate-2; STEAP-2;
KW	chromosome 7q21; cancer; prostate; colon; bladder; lung; ovarian;
KX	pancreatic.
XX	XX
OS	Homo sapiens.
XX	XX
FH	Key
FT	Region
FT	Region
FT	Region
FT	Domain
FT	Domain
FT	Region
FT	Region
FT	Region
FT	Region
FT	Domain
FT	Domain
FT	Domain
FT	Domain
FT	Region
FT	Region
FT	Region
FT	Domain
FT	Domain
PN	WO200140276-A2.
XX	07-JUN-2001.
PD	

Location/Qualifiers

/label= HLA-A2_binding_peptide #5

/label= Immunogenic_peptide #1

/label= Transmembrane_domain #1

/label= Transmembrane_domain #1

/label= Transmembrane_domain #2

/label= Transmembrane_domain #2

/label= Transmembrane_domain #3

/label= Transmembrane_domain #3

/label= HLA-A2_binding_peptide #4

/label= HLA-A2_binding_peptide #4

/label= Immunogenic_peptide #2

/label= Transmembrane_domain #4

/label= Transmembrane_domain #5

/label= HLA-A2_binding_peptide #2

/label= 450

/label= Transmembrane_domain #6

XX 06-DEC-2000; 2000WO-US033040.
PF 06-DEC-1999; 99US-00455486.
PR (UROG-) UROGENESYS INC.
XX Afar DEH, Hubert RS, Raitano AB, Saffran DC, Mitchell SC;
PI Faris M, Jakobovits A;
XX WPI; 2001-367804/38.
DR N-PSDB; AAD07072.
XX
PT New STEAP (six transmembrane epithelial antigen of the prostate)
PT proteins, expressed in human cancers, useful for detecting and treating
PT cancer.
XX
PS Claim 1; Fig 9A-9C; 187bp; English.
XX
CC The present sequence is human six transmembrane epithelial antigen of the
CC prostate (STEAP)-2 protein. STEAP is a member of cell surface serpentine
CC transmembrane antigens. STEAP-2 gene is located on chromosome 7q21 and is
CC used in gene therapy. Inhibiting the development or progression of a
CC cancer (eg. prostate, colon, bladder, lung, ovarian and pancreatic)
CC expressing STEAP or inhibiting growth or killing cells expressing STEAP
CC in a patient, comprises administering a vaccine composition to the
CC patient. Treating a patient with a cancer that expresses STEAP, or
CC inhibiting growth or killing cells expressing STEAP, comprises
CC administering to the patient a vector encoding single chain monoclonal
CC antibody that comprises the variable domains of the heavy and light
CC chains of the monoclonal antibody that specifically binds to STEAP, such
CC that the vector delivers the single chain monoclonal antibody coding
CC sequence to the cancer cells and the encoded single chain monoclonal
CC antibody is expressed intracellularly. Note: This sequence is stated to
CC be the same as that being shown as SEQ ID NO:8 (AAE02841) in sequence
CC listing of the specification. However both the sequences differ at
CC several positions
XX
XX Sequence 454 AA;
SQ
Query Match 100.0%; Score 2351; DB 4; Length 454;
Best Local Similarity 100.0%; Pred. No. 5e-245;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 421 EYYRYTPPNFVLALVLPSTIVIDLQLCRYPD 454
|||||
RESULT 2
AAU10188
ID AAU10188 standard; protein; 454 AA.
XX
XX AAU10188;
AC
XX 16-JAN-2002 (first entry)
DT
XX Human ORF2 of Six-Transmembrane Protein of Prostate 1.
DE
XX Human; Six-Transmembrane Protein of Prostate 1; STMP1; prostate cancer;
KW benign prostatic hyperplasia; acute prostatitis; testicular cancer;
KW cryptorchidism; testicular disorder; proliferative disorder; lymphoma;
KW leukaemia; melanoma; ovarian cancer; breast cancer; pancreatic cancer;
KW liver cancer; lung cancer; cytostatic; ORF2.
KM
XX Homo sapiens.
OS
XX WO200172962-A2.
PN
XX 04-OCT-2001.
PD
XX 23-MAR-2001; 2001WO-US009410.
PF
XX 24-MAR-2000; 2000US-0191929P.
PR
XX (SAAT/) SAATCIOGLU F.
PA
XX Saatioglu F;
PI
XX WPI; 2001-662926/76.
DR N-PSDB; AAS15810, AAS15811.
DR
XX
PT New polynucleotide for the diagnosis, prevention and treatment for
PT prostate and testis disorders, particularly prostate cancer, comprises
PT prostate-specific or testis-specific nucleic acids.
PT
XX
XX Claim 1; Fig 4H; 114pp; English.
PS
XX The invention relates to substantially pure prostate-specific or testis-
XX specific polypeptides and the nucleic acids encoding them. Also included
XX are vectors and host cells expressing the proteins, a transgenic animal
XX expressing the protein, antibodies against the proteins, probes for
XX detecting the nucleic acids, antisense molecules for the nucleic acids
XX and methods of isolating modulators of the proteins. Compounds that
XX modulate the prostate specific or testis specific polypeptide are useful
XX to diagnose, prevent or treat disorders of the testis or prostate
XX particularly prostate cancer, benign prostatic hyperplasia, acute
XX prostatitis, testicular cancer, cryptorchidism, undescended, retractile,
XX ascending or vanished testis. Other proliferative disorders for which the
XX modulators may be used include lymphoma, leukaemia, melanoma, ovarian
XX cancer, breast cancer, pancreatic cancer, liver cancer and lung cancer.
XX The present sequence is prostate specific protein, Six-Transmembrane
XX protein of prostate 1, STMP1, ORF2
XX
XX Sequence 454 AA;
SQ
Query Match 100.0%; Score 2351; DB 4; Length 454;
Best Local Similarity 100.0%; Pred. No. 5e-245;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 RINQYPESNAEYLAFLPDSLIVKGFNVSAWALQIGPKDASROYICSNNIQARQVIE 180
CC |||||
CC RINQYPESNAEYLAFLPDSLIVKGFNVSAWALQIGPKDASROYICSNNIQARQVIE 180
Db 121 RINQYPESNAEYLAFLPDSLIVKGFNVSAWALQIGPKDASROYICSNNIQARQVIE 180
QY 181 LARQINFIPIDLGSLSSAREIENLPRLFTLMRGPPVVAISLATFFFLYSFVRDVIHPYA 240
CC |||||
CC LARQINFIPIDLGSLSSAREIENLPRLFTLMRGPPVVAISLATFFFLYSFVRDVIHPYA 240
Db 181 LARQINFIPIDLGSLSSAREIENLPRLFTLMRGPPVVAISLATFFFLYSFVRDVIHPYA 240
QY 241 RNOQSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTXYRRFPFWLETWLQ 300
CC |||||
CC RNOQSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTXYRRFPFWLETWLQ 300
Db 241 RNOQSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTXYRRFPFWLETWLQ 300
QY 301 CRKQGLISFFPAMVHVAYSLCLPMRRSERYLFLMAYQQOVHANIENSMNEEYWRRIEMX 360
CC |||||
CC CRKQGLISFFPAMVHVAYSLCLPMRRSERYLFLMAYQQOVHANIENSMNEEYWRRIEMX 360
Db 301 CRKQGLISFFPAMVHVAYSLCLPMRRSERYLFLMAYQQOVHANIENSMNEEYWRRIEMX 360
QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHYLIYGMKRAFE 420
CC |||||
CC ISFGIMSLGLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHYLIYGMKRAFE 420
Db 361 ISFGIMSLGLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHYLIYGMKRAFE 420
QY 421 EEEYRFPYTPPNFVLALVLPISIVILDLQLCRYPD 454
CC |||||
CC EEEYRFPYTPPNFVLALVLPISIVILDLQLCRYPD 454
Db 421 EEEYRFPYTPPNFVLALVLPISIVILDLQLCRYPD 454

RESULT 3

ABG61933
ID ABG61933 standard; protein; 454 AA.

XX AC ABG61933;

XX DT 15-AUG-2002 (first entry)

XX DE Prostate cancer-associated protein #134.

XX KM Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.

XX OS Mammalia.

XX PN WO200230268-A2.

XX PD 18-APR-2002.

XX PF 12-OCT-2001; 2001WO-US032045.

XX PR 13-OCT-2000; 2000US-00687576.

XX PR 08-DEC-2000; 2000US-00733288.

XX PR 08-DEC-2000; 2000US-00733742.

XX PR 24-JAN-2001; 2001US-0263957P.

XX PR 16-MAR-2001; 2001US-0276791P.

XX PR 16-MAR-2001; 2001US-0276888P.

XX PR 06-APR-2001; 2001US-0281922P.

XX PR 24-APR-2001; 2001US-0286214P.

XX PR 30-APR-2001; 2001US-00847046.

XX PR 04-MAY-2001; 2001US-0288589P.

XX PA (EOSB-) EOS BIOTECHNOLOGY INC.

XX PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;

XX PS WPI; 2002-471335/50.

XX DR N-PSDB; ABK92252.

XX PT Detecting a prostate cancer-associated transcript in a cell in a patient,
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
PT by determining if prostate cancer-associated genes are expressed in a
PT prostate tissue.

XX PS Claim 27; Page 416; 436pp; English.

XX CC The present invention relates to methods of detecting a prostate cancer-
CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with prostate cancer-

CC associated polynucleotides (designated PC genes) that selectively
CC hybridise to a sequence that is at least 80% identical to them. The
CC prostate cancer-associated polynucleotide sequences are differentially
CC expressed in prostate tumour tissue or in prostate cancer and are derived
CC from the tissues of various organisms such as humans or other mammals
CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
CC associated genes are useful for diagnosing or treating prostate cancer,
CC as well as for identifying modulators of prostate cancer or agents that
CC inhibit prostate cancer. The nucleic acid sequences are particularly
CC useful in gene therapy, as a vaccine or in antisense applications.
CC ABG61800-ABG61944 represent prostate cancer-associated proteins

XX SQ Sequence 454 AA;

Query Match 100.0%; Score 2351; DB 5; Length 454;
Best Local Similarity 100.0%; Pred. No. 5e-245;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPENGINGIKDARKVTVGVIGSGDFAKSLTIRLCGYHVVIGS 60
Db 1 MESISMGSPKSLSETCLPENGINGIKDARKVTVGVIGSGDFAKSLTIRLCGYHVVIGS 60

QY 61 RNDKFASEFPFPHVVDVTHHEDALTKNTIIFVAIHREHYTSLMDLRHLLVGKILIDVSNM 120
Db 61 RNDKFASEFPFPHVVDVTHHEDALTKNTIIFVAIHREHYTSLMDLRHLLVGKILIDVSNM 120

QY 121 RINQYPESNAEYLAFLPDSLIVKGFNVSAWALQIGPKDASROYICSNNIQARQVIE 180
Db 121 RINQYPESNAEYLAFLPDSLIVKGFNVSAWALQIGPKDASROYICSNNIQARQVIE 180

QY 181 LARQINFIPIDLGSLSSAREIENLPRLFTLMRGPPVVAISLATFFFLYSFVRDVIHPYA 240
Db 181 LARQINFIPIDLGSLSSAREIENLPRLFTLMRGPPVVAISLATFFFLYSFVRDVIHPYA 240

QY 241 RNOQSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTXYRRFPFWLETWLQ 300
Db 241 RNOQSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTXYRRFPFWLETWLQ 300

QY 301 CRKQGLISFFPAMVHVAYSLCLPMRRSERYLFLMAYQQOVHANIENSMNEEYWRRIEMX 360
Db 301 CRKQGLISFFPAMVHVAYSLCLPMRRSERYLFLMAYQQOVHANIENSMNEEYWRRIEMX 360

QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHYLIYGMKRAFE 420
Db 361 ISFGIMSLGLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHYLIYGMKRAFE 420

QY 421 EEEYRFPYTPPNFVLALVLPISIVILDLQLCRYPD 454
Db 421 EEEYRFPYTPPNFVLALVLPISIVILDLQLCRYPD 454

RESULT 4

AAU76538
ID AAU76538 standard; protein; 454 AA.

XX AC AAU76538;

XX DT 05-JUN-2002 (first entry)

XX DE Tumour-associated antigenic target protein, TAT138.

XX KM TAT138; Tumour-associated Antigenic Target; tumour; breast cancer;

XX KM colorectal cancer; lung cancer; ovarian cancer;

XX KM central nervous system cancer; liver cancer; bladder cancer; melanoma;

XX KM pancreatic cancer; leukaemia; gene therapy.

XX OS Homo sapiens.

XX FH Key

XX FT Modified-site 8..13
XX FT Modified-site /note= "N-myristoylation site"
XX FT Modified-site 24..29

FT /note= "N-myristoylation site"
 FT Region 29.32
 FT /note= "cAMP- and cGMP-dependent protein kinase
 FT phosphorylation site"
 FT Modified-site 34.39
 FT /note= "N-myristoylation site"
 FT Modified-site 193.198
 FT /note= "N-myristoylation site"
 FT Domain 210.230
 FT /note= "Transmembrane domain"
 FT Modified-site 256.259
 FT /note= "Asn is N-glycosylated"
 FT Domain 257.277
 FT /note= "Transmembrane domain"
 FT Domain 259.379
 FT /note= "Transmembrane domain"
 FT Modified-site 274.279
 FT /note= "N-myristoylation site"
 FT Domain 299.319
 FT /note= "Transmembrane domain"
 FT Domain 393.413
 FT /note= "Transmembrane domain"
 FT Modified-site 416.424
 FT /note= "Tyrosine kinase phosphorylation site"
 FT Domain 428.448
 FT /note= "Transmembrane domain"
 PN WO200216429-A2.
 XX 28-FEB-2002.
 XX 22-JUN-2001; 2001WO-US020118.
 XX 24-AUG-2000; 2000WO-US023328.
 PR 26-SEP-2000; 2000US-0235451P.
 PR 01-DEC-2000; 2000WO-US032678.
 PR 28-FEB-2001; 2001WO-US006520.
 PR 01-MAR-2001; 2001WO-US006666.
 XX (GETH) GENENTECH INC.
 XX Goddard A, Godowski PJ, Gurney AL, Hillan KJ, Polakis P, Smith V;
 PI Wood WJ, Wu TD, Zhang Z;
 XX WPI; 2002-280917/32.
 DR N-PSDB; ABK11093.
 XX Novel isolated tumor-associated antigenic target polypeptides which are
 PT useful as targets for cancer therapy and diagnosis in mammals.
 XX Claim 12; Fig 10; 121pp; English.
 CC The invention relates to an isolated tumour-associated antigenic target
 CC polypeptide (TAT) (I), specifically TAT134-TAT138 polypeptides, and the
 CC polynucleotides (II) encoding them. (II) is useful for diagnosing the
 CC presence of a tumour in a mammal, where the level of expression of (II)
 CC is indicative of the presence of tumour in the mammal from which the test
 CC sample was obtained. Antibody to (I) is useful for killing a cancer cell
 CC (e.g. breast cancer cell, a colorectal cancer cell, a lung cancer cell,
 CC an ovarian cancer cell, a central nervous system (CNS) cancer cell, a
 CC liver cancer cell, a bladder cancer cell, a pancreatic cancer cell, a
 CC melanoma cell or a leukemia cell) that expresses (I). Oligonucleotides
 CC hybridising to (II) are useful as diagnostic probes, antisense
 CC oligonucleotide probes or for encoding fragments of full length TAT
 CC polypeptide. (II) is also useful in chromosome and gene mapping and in
 CC the generation of antisense RNA and DNA probes, for constructing
 CC hybridisation probes for mapping the gene encoding TAT and for genetic
 CC analysis of individuals with genetic disorders. (II) is also useful for
 CC generating either transgenic animals or knockout animals, and in gene
 CC therapy. The TAT polypeptides and nucleic acids may also be used for
 CC tissue typing and the TAT polypeptides are useful for screening compounds
 CC that mimic the TAT polypeptide (agonist) or prevent the effect of TAT
 CC polypeptide (antagonist). The antibody is useful for staging TAT

CC polypeptide-expressing cancers, purifying or immunoprecipitating TAT
 CC polypeptide from cells, for detection and quantitation of TAT polypeptide
 CC in vitro, e.g., in an enzyme linked immunosorbent assay (ELISA) or
 CC Western blot. The antibodies are also useful for treating a TAT-
 CC expressing cancer or alleviating one or more symptoms of cancer in a
 CC mammal. The present sequence represents the amino acid sequence of TAT138
 CC
 SQ Sequence 454 AA;
 Query Match 100.0%; Score 2351; DB 5; Length 454;
 Best Local Similarity 100.0%; Pred. No. 5e-245;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MESISMGMSPKSLSEICLPNGINGIKDKAKVTGVIGSGDFAKSLTIRLRGYNVIGS 60
 DB 1 MESISMGMSPKSLSEICLPNGINGIKDKAKVTGVIGSGDFAKSLTIRLRGYNVIGS 60
 QY 61 RNPKEASEFPFHVVDVTHHEDALTKNTIFVAIHREHYTSLMDLRHLLVGKILIDVSNM 120
 DB 61 RNPKEASEFPFHVVDVTHHEDALTKNTIFVAIHREHYTSLMDLRHLLVGKILIDVSNM 120
 QY 121 RINQYPSNAEYLALEPDSLIVKGFNVSAWALQGPCKDASQYVICSNNIQARQVIE 180
 DB 121 RINQYPSNAEYLALEPDSLIVKGFNVSAWALQGPCKDASQYVICSNNIQARQVIE 180
 QY 181 LARQNFIPIDLGSLSSAREIENPLRLFTLMRGPVVVAISLATPFLYSFVRDVIHPYA 240
 DB 181 LARQNFIPIDLGSLSSAREIENPLRLFTLMRGPVVVAISLATPFLYSFVRDVIHPYA 240
 QY 241 RNOQSDPYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTYKRRPPLMTWLO 300
 DB 241 RNOQSDPYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTYKRRPPLMTWLO 300
 QY 301 CRKQIGLISFPFAMVHAYSLCLPMRSEERYLFLNAYQOVHANENSWNEBEVRIEY 360
 DB 301 CRKQIGLISFPFAMVHAYSLCLPMRSEERYLFLNAYQOVHANENSWNEBEVRIEY 360
 QY 361 ISFGIMSLGLSLAVTSIPSVSNALNMRSEFSIOSTLGVALLISTFHVLIYGMKRAPE 420
 DB 361 ISFGIMSLGLSLAVTSIPSVSNALNMRSEFSIOSTLGVALLISTFHVLIYGMKRAPE 420
 QY 421 EBYRFFYTPPNFVLAIVLPISVIVLDLQLCRYPD 454
 DB 421 EBYRFFYTPPNFVLAIVLPISVIVLDLQLCRYPD 454
 RESULT 5
 AAU80190
 ID AAU80190 standard; protein; 454 AA.
 XX AAU80190;
 AC AAU80190;
 XX 15-JUL-2002 (first entry)
 DE Human PUMPCn protein, PRO23203.
 XX Human; protein upregulated in metastatic prostate cancer; immunogen;
 KW PUMPCn; PRO23203; prostate cancer; cytostatic; gene therapy; transgenic;
 KW androgen independent prostate cancer; DNA microarray.
 XX Homo sapiens.
 OS Homo sapiens.
 PN WO200226822-A2.
 XX 04-APR-2002.
 PD 26-SEP-2001; 2001WO-US030290.
 PF 26-SEP-2000; 2000US-0235451P.
 PR 26-SEP-2000; 2000US-0235451P.
 XX (GETH) GENENTECH INC.
 PA Devaux B, Eberhard D, Goddard A, Godowski PJ, Grimaldi JC;
 PI

PI Hillan KJ, Watanabe CK, Wood WI, Yansura DG, Zhang Z;
XX WPI: 2002-383270/41.
DR N-PSDB; ABK50391.
XX
PT New polypeptide termed protein upregulated in metastatic prostate cancer
PT and encoding polynucleotides, useful for identifying polypeptide
PT antagonists for treating prostate cancer.
XX
PS Claim 23; Fig 2; 137pp; English.
XX

CC The invention relates to an isolated human protein upregulated in
CC metastatic prostate cancer (PUMPCn) PRO23203 polypeptide, a sequence 80%
CC identical to PRO23203 and the sequence as encoded by cDNA insert of the
CC vector deposited as ATCC Deposit No. PTA-2513 (DNA185171- 2994) on
CC 26/9/2000. Also included are the polynucleotide encoding the protein (or
CC a DNA sequence 80% identical to the polynucleotide and one that
CC hybridises to complement of the polynucleotide), a vector comprising the
CC polynucleotide, a polynucleotide deposited with ATCC under accession
CC number PTA-2513 (DNA185171-2994), a host cell comprising the vector,
CC preparation of PRO23203, a chimeric molecule comprising PRO23203 fused to
CC a heterologous amino acid sequence, an anti-PRO23203 antibody, an
CC agonist/antagonist of PRO23203, and diagnosing the presence of prostate
CC cancer in a mammal by: (a) contacting a microarray diagnostic with a
CC DNA185171-2994 probe, detecting and quantifying hybridisation of
CC DNA185171-2994 probe in prostate cancer tissue compared with normal
CC tissue and determining if DNA185171-2994 is overexpressed; or (b)
CC contacting a tissue of the mammal with an anti-PRO23203 antibody and
CC detecting the binding of the antibody to a component of the tissue, where
CC binding is indicative of the presence of prostate cancer in the mammal.
CC The antibody is useful for treating prostate cancer in mammal which is
CC androgen independent prostate cancer, that has metastasised to another
CC portion of the body, where the antibody is not conjugated with a
CC cytotoxic agent and the method further comprises administering a
CC chemotherapeutic agent to the mammal and for diagnosing the presence of
CC prostate cancer in an mammal. The PRO23203 polynucleotide is useful for
CC generating either transgenic animals or knock out animals which in turn
CC are useful in the development and screening of therapeutically useful
CC reagents or for use in gene therapy and for chromosome identification.
CC PRO23203 and polynucleotide may also be used for tissue typing and
CC PRO23203 may also be employed as a therapeutic agents and for screening
CC compounds to identify those that mimic the PRO23203 polypeptides
CC (agonists) or prevent the effect of the PRO23203 polypeptide
CC (antagonists). PRO23203 is further useful for the affinity purification
CC of PRO23203 from recombinant cell culture or natural sources. The present
CC sequence represents PRO23203
XX
XX Sequence 454 AA;
SQ

Query Match 100.0%; Score 2351; DB 5; Length 454;
Best Local Similarity 100.0%; Pred. No. 5e-245;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHVG 60
Db 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHVG 60
QY 61 RNPKEASEFPFHVDVTHHEDALTNTNIFVAIHREHYTSLWDLRHLVGKILIDVSNM 120
Db 61 RNPKEASEFPFHVDVTHHEDALTNTNIFVAIHREHYTSLWDLRHLVGKILIDVSNM 120
QY 121 RINQYPSNAEYLAFLPDSILVKGFNVSAMALQGPKDASRQVYICSNNOARQOVIE 180
Db 121 RINQYPSNAEYLAFLPDSILVKGFNVSAMALQGPKDASRQVYICSNNOARQOVIE 180
QY 181 IARQINFIPIDLGLSSAREIENLPLRLFTLWRGPVVAISLATFFFLYSFVDVHPYA 240
Db 181 IARQINFIPIDLGLSSAREIENLPLRLFTLWRGPVVAISLATFFFLYSFVDVHPYA 240
QY 241 RNOQSDFYKIPIEIVNKTLPVATLTLVLVLAAGLAAAYQLYGTQYRRFPWLETWLQ 300
Db 241 RNOQSDFYKIPIEIVNKTLPVATLTLVLVLAAGLAAAYQLYGTQYRRFPWLETWLQ 300

QY 301 CRKQLGLSFFPAMVAVYSLCLPMRRSERVFLMAYQOVHANINENWNEEVRRIEM 360
Db 301 CRKQLGLSFFPAMVAVYSLCLPMRRSERVFLMAYQOVHANINENWNEEVRRIEM 360
QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRERSFIQSTLGYVALLISTHVLIIYWKRAFE 420
Db 361 ISFGIMSLGLSLAVTSIPSVSNALNWRERSFIQSTLGYVALLISTHVLIIYWKRAFE 420
QY 421 EBYRYFTPPNFVLAIVLPSIVILDLQLCRYPD 454
Db 421 EBYRYFTPPNFVLAIVLPSIVILDLQLCRYPD 454

RESULT 6
ABU08893
ID ABU08893 standard; protein; 454 AA.
XX
AC ABU08893;
XX
DT 09-OCT-2003 (first entry)
XX
DE Tumour-associated antigenic target 138 (TAT138).
XX
KW Human; tumour-associated antigenic target polypeptide; TAT; tumour;
KW gene therapy; cytostatic; tissue typing; prostate tumour; cancer.
XX
OS Homo sapiens.
XX
PN US2003060612-A1.
XX
PD 27-MAR-2003.
XX
PF 22-JUN-2001; 2001US-00888257.
XX
PR 28-OCT-1997; 97US-0063540P.
PR 17-JUN-1998; 98US-0089653P.
PR 10-SEP-1998; 98US-0099792P.
PR 08-OCT-1998; 98US-0103678P.
PR 02-JUN-1999; 99WO-US012252.
PR 01-SEP-1999; 99WO-US020111.
PR 18-FEB-2000; 2000WO-US004342.
PR 02-MAR-2000; 2000WO-US005841.
PR 30-MAR-2000; 2000WO-US008439.
PR 24-AUG-2000; 2000WO-US023328.
PR 26-SEP-2000; 2000US-0235451P.
PR 01-DEC-2000; 2000WO-US032678.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
XX
PA (GETH) GENENTECH INC.
XX
PI Goddard A, Godowski PJ, Gurney AL, Hillan KJ, Polakis P, Smith V;
PI Wood WI, Wu TD, Zhang Z;
XX
XX WPI: 2003-576355/54.
DR N-PSDB; ACD25893.
XX
PT New nucleic acid, useful for the manufacture of a medicament for
PT diagnosing or treating tumor in a mammal.
XX
PS Claim 12; Fig 10; 71pp; English.
XX
CC The invention discloses human nucleic acids encoding tumour-associated
CC antigenic target (TAT) polypeptides, with or without their associated
CC signal peptide. Also disclosed is an antibody that specifically binds to
CC the TAT polypeptides, a method for detecting the presence of a tumour in
CC a mammal and a method for killing a cancer cell expressing the TAT
CC polypeptide. The nucleotide sequences are useful in preparing TAT
CC polypeptides by recombinant techniques and in gene therapy (e.g. for
CC replacement of defective gene). The TAT polypeptides are useful as
CC therapeutic agents and for detecting the presence, prevention and/or
CC treatment of a tumour, such as colon, breast or prostate tumour. The TAT
CC polypeptides and nucleic acids may also be used diagnostically for tissue

CC typing. The sequence presented is the TAT138 polypeptide of the invention
XX Sequence 454 AA;

Query Match 100.0%; Score 2351; DB 7; Length 454;
Best Local Similarity 100.0%; Pred. No. 5e-245;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIIGSGDPAKSLTIRLCGYHVVIGS 60
DB 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIIGSGDPAKSLTIRLCGYHVVIGS 60
QY 61 RNPKEASEFPFHVDVTHHEDALTKNTIIFVAIHREHYTSLMDLRHLLVSKILIDVSNM 120
DB 61 RNPKEASEFPFHVDVTHHEDALTKNTIIFVAIHREHYTSLMDLRHLLVSKILIDVSNM 120
QY 121 RINQYPSNAEYLALSPDLSLVKGFNVVSAMALQGPKASRQVYICSNNIQARQVIE 180
DB 121 RINQYPSNAEYLALSPDLSLVKGFNVVSAMALQGPKASRQVYICSNNIQARQVIE 180
QY 181 LARQNFIPIDIGSLSSAREIENLPLRLFTLMRGPVVVAISLATFFFLYSFVRDVIHPYA 240
DB 181 LARQNFIPIDIGSLSSAREIENLPLRLFTLMRGPVVVAISLATFFFLYSFVRDVIHPYA 240
QY 241 RNQSDFYKIPIEIVNKTLPVIAITLLSLVYLAGLLAAAYQLYGTYRRPFPWLETWLQ 300
DB 241 RNQSDFYKIPIEIVNKTLPVIAITLLSLVYLAGLLAAAYQLYGTYRRPFPWLETWLQ 300
QY 301 CRKQLGLSFFPAMVHVAISLCLPMRRESERYLFNMAYQQVHANINENSWNEEVRRIEM 360
DB 301 CRKQLGLSFFPAMVHVAISLCLPMRRESERYLFNMAYQQVHANINENSWNEEVRRIEM 360
QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRRESFIQSTLGYVALLISTFHVLIYGWKRAFE 420
DB 361 ISFGIMSLGLSLAVTSIPSVSNALNWRRESFIQSTLGYVALLISTFHVLIYGWKRAFE 420
QY 421 EEYRFRYTPPNFVLALVLPISIVILDLLQCRYPD 454
DB 421 EEYRFRYTPPNFVLALVLPISIVILDLLQCRYPD 454

RESULT 7
ADB65001
ID ADB65001 standard; protein; 454 AA.
XX ADB65001;
AC ADB65001;
XX 04-DEC-2003 (first entry)
DT 04-DEC-2003 (first entry)
XX Human protein encoded by clone PROST20168600.
DE Human protein encoded by clone PROST20168600.
XX Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;
KW cell regeneration; membrane protein; signal transduction-related protein;
KW transcription-related protein; osteoporosis; neurological disease;
KW cancer; tumour.
XX Homo sapiens.
OS Homo sapiens.
XX EPI308459-A2.
PN EPI308459-A2.
XX 07-MAY-2003.
PD 07-MAY-2003.
XX 28-MAR-2002; 2002EP-00007401.
PF 28-MAR-2002; 2002EP-00007401.
XX 05-NOV-2001; 2001JP-00379298.
PR 25-JAN-2002; 2002US-00350978.
XX (HELI-) HELIX RES INST.
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX (HELI-) HELIX RES INST.
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

XX WPI: 2003-450961/43.
DR N-PSDB; ADB63031.
DR N-PSDB; ADB63031.

PT New polynucleotides and polypeptides, useful for developing a diagnostic
PT marker or medicines for regulation of their expression and activity, or
PT as targets of gene therapy.

XX Claim 1; Page; 222pp; English.

CC The invention discloses a polynucleotide comprising a sequence selected
CC from 1970 fully defined nucleotide sequences which encode novel
CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
CC or its partial peptide, an antibody binding to the polypeptide or peptide
CC of the polynucleotide, immunologically assaying the polypeptide or peptide
CC peptide of the polynucleotide by contacting the polypeptide or peptide
CC with the antibody of the encoded protein, and observing the binding
CC between the two, a transformant carrying the polynucleotide in an
CC expressible manner and an antisense polynucleotide. The oligonucleotide
CC is useful as a primer for synthesizing the polynucleotide, or as a probe
CC for detecting the polynucleotide. The polynucleotides and encoded
CC proteins are useful as pharmaceutical agents and many disease-related
CC genes may be included in them, for developing a diagnostic marker or
CC medicines for regulation of their expression and activity, or as targets
CC of gene therapy. The genes are involved in tissue and/or cell
CC regeneration. Membrane proteins, signal transduction-related proteins,
CC transcription-related proteins, disease-related proteins and genes
CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
CC neurological diseases, cancer, tumours. The cDNA may be used to regulate
CC the activity or expression of the encoded protein to treat diseases. The
CC sequence presented is a protein of the invention. Note: Some of the
CC sequence data for this patent is not represented in the printed
CC specification, but is based on sequence information supplied by the
CC European Patent Office.

XX Sequence 454 AA;

Query Match 99.9%; Score 2348; DB 7; Length 454;
Best Local Similarity 99.8%; Pred. No. 1.1e-244;
Matches 453; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIIGSGDPAKSLTIRLCGYHVVIGS 60
DB 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIIGSGDPAKSLTIRLCGYHVVIGS 60
QY 61 RNPKEASEFPFHVDVTHHEDALTKNTIIFVAIHREHYTSLMDLRHLLVSKILIDVSNM 120
DB 61 RNPKEASEFPFHVDVTHHEDALTKNTIIFVAIHREHYTSLMDLRHLLVSKILIDVSNM 120
QY 121 RINQYPSNAEYLALSPDLSLVKGFNVVSAMALQGPKASRQVYICSNNIQARQVIE 180
DB 121 RINQYPSNAEYLALSPDLSLVKGFNVVSAMALQGPKASRQVYICSNNIQARQVIE 180
QY 181 LARQNFIPIDIGSLSSAREIENLPLRLFTLMRGPVVVAISLATFFFLYSFVRDVIHPYA 240
DB 181 LARQNFIPIDIGSLSSAREIENLPLRLFTLMRGPVVVAISLATFFFLYSFVRDVIHPYA 240
QY 241 RNQSDFYKIPIEIVNKTLPVIAITLLSLVYLAGLLAAAYQLYGTYRRPFPWLETWLQ 300
DB 241 RNQSDFYKIPIEIVNKTLPVIAITLLSLVYLAGLLAAAYQLYGTYRRPFPWLETWLQ 300
QY 301 CRKQLGLSFFPAMVHVAISLCLPMRRESERYLFNMAYQQVHANINENSWNEEVRRIEM 360
DB 301 CRKQLGLSFFPAMVHVAISLCLPMRRESERYLFNMAYQQVHANINENSWNEEVRRIEM 360
QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRRESFIQSTLGYVALLISTFHVLIYGWKRAFE 420
DB 361 ISFGIMSLGLSLAVTSIPSVSNALNWRRESFIQSTLGYVALLISTFHVLIYGWKRAFE 420
QY 421 EEYRFRYTPPNFVLALVLPISIVILDLLQCRYPD 454
DB 421 EEYRFRYTPPNFVLALVLPISIVILDLLQCRYPD 454

RESULT 8
AAE28951
ID AAE28951 standard; protein; 490 AA.
XX
AC AAE28951;
XX
DT 27-JAN-2003 (first entry)
XX
DE Human STEAPRP protein.
XX
KM Human; STEAPRP, proliferative disorder; prostate hyperplasia; cancer;
KM six-transmembrane epithelial antigen of the prostate related protein;
KM gene therapy; cytostatic.
XX
OS Homo sapiens.
XX
PN WO200272596-A1.
XX
PD 19-SEP-2002.
XX
PF 07-MAR-2002; 2002WO-US007053.
XX
PR 09-MAR-2001; 2001US-00802520.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Lal PG, Faris M, Chen H, Ison CH;
XX
DR WPI: 2002-723311/78.
DR N-PSDB; AAD46429.
XX
PT New six-transmembrane epithelial antigen of the prostate-related protein
PT and nucleic acids, useful for diagnosing and treating prostate cell
PT proliferative disorders, particularly prostate hyperplasia and prostate
PT cancer.
XX
PS Claim 13; Page 54-55; 62pp; English.
XX
CC The invention relates to six-transmembrane epithelial antigen of the
CC prostate related protein (STEAPRP) and its corresponding nucleic acid.
CC STEAPRP DNA is used to diagnose and to treat prostate cell proliferative
CC disorders, particularly prostate hyperplasia and prostate cancer, and for
CC screening a library of molecules of compounds for specific binding
CC affinity. It is also used in gene therapy. STEAPRP is used to screen a
CC subject sample for antibodies, which specifically binds the protein, and
CC to prepare and purify a protein. The present sequence is human STEAPRP
CC protein
CC
SQ Sequence 490 AA;

Query Match 97.6%; Score 2294; DB 5; Length 490;
Best Local Similarity 100.0%; Pred. No. 8.4e-239;
Matches 444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 241 RNOQSDFYKIPIEIVNKTLPVAILTLISLVYLAGLLAAAYQLYGTYRRFPFWLFTWLQ 300
QY 301 CRKQIGLSFFEFAMVHVAVSLCLPMRSEERYLFNMAVQOVHANIENSWNEEYVRIEMV 360
DB 301 CRKQIGLSFFEFAMVHVAVSLCLPMRSEERYLFNMAVQOVHANIENSWNEEYVRIEMV 360
QY 361 ISFGIMSLGLSLLAVTSIPSVSNALNWRFEFSFIQSTLGVALLISTFHVLIYGWKRAFE 420
DB 361 ISFGIMSLGLSLLAVTSIPSVSNALNWRFEFSFIQSTLGVALLISTFHVLIYGWKRAFE 420
QY 421 EBYRFRYTPPNFVLAIVLPSIVIL 444
DB 421 EBYRFRYTPPNFVLAIVLPSIVIL 444

RESULT 9
ABU63312
ID ABU63312 standard; protein; 490 AA.
XX
AC ABU63312;
XX
DT 18-SEP-2003 (first entry)
XX
DE Human transmembrane protein TMPL.
XX
KM Human; prostate tumour; lung tumour; TMPL; cancer.
XX
OS Homo sapiens.
XX
PN US2003064397-A1.
XX
PD 03-APR-2003.
XX
PF 24-JUL-2002; 2002US-00205267.
XX
PR 22-MAY-1998; 98US-00083521.
PR 16-SEP-1999; 99US-00397558.
PR 09-MAR-2001; 2001US-00802520.
PR 26-SEP-2001; 2001US-00963896.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Spancake KM, Rickert PK, Lal PG, Ison CH;
XX
DR WPI: 2003-540783/51.
DR N-PSDB; ACD27642.
XX
PT New transmembrane protein differentially expressed in prostate and lung
PT tumors, useful for diagnosing, staging, treating or monitoring
PT progression or treatment of cancer, preferably lung or prostate cancer.
XX
PS Claim 1; Fig 1; 49pp; English.
XX
CC The invention relates to a purified human protein comprising a
CC polypeptide having a fully defined transmembrane protein differentially
CC expressed in prostate and lung tumour (TMPL) sequence. The protein is
CC useful for detecting expression of the protein in a sample, where the
CC sample is lung or prostate tissue sample and the protein is
CC differentially expressed when compared with a standard and is diagnostic
CC of a lung or prostate cancer. The protein is useful to identify an
CC antibody that specifically binds the protein. The protein is useful for
CC screening several molecules and compounds to identify at least one
CC ligand. The protein is also useful for preparing a monoclonal antibody using
CC standard animal immunisation/hybridoma methods. A TMPL specific antibody
CC is useful to detect expression of the protein in a lung or prostate
CC tissue sample and the method is useful as a diagnostic for lung or
CC prostate cancer. A TMPL specific antibody is also useful for
CC immunopurifying a protein, for treating a prostate cancer and for
CC delivering a therapeutic agent to a cancer, preferably prostate cancer.
CC The protein is useful for diagnosing, staging, treating or monitoring
CC progression or treatment of cancer, preferably lung or prostate cancer.
CC The present sequence represents the amino acid sequence of the human

transmembrane protein differentially expressed in prostate and lung
tumour TMPL
Sequence 490 AA;

Query Match 97.6%; Score 2294; DB 7; Length 490;
Best Local Similarity 100.0%; Pred. No. 8.4e-239;
Matches 444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPNGINGIKARKVTGVIGSGDPAKSLTIRLRGCVHVGIS 60
DB 1 MESISMGSPKSLSETCLPNGINGIKARKVTGVIGSGDPAKSLTIRLRGCVHVGIS 60
QY 61 RNPKEASEFPFHVDTTHEDALTKNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
DB 61 RNPKEASEFPFHVDTTHEDALTKNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
QY 121 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQGPXKASRQVYICSNNIQARQVIE 180
DB 121 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQGPXKASRQVYICSNNIQARQVIE 180
QY 181 LARQNFIPIDGLSSAREIENLPLRFTLWRGPVVAISLATFFFLYSFVRDVTHPYA 240
DB 181 LARQNFIPIDGLSSAREIENLPLRFTLWRGPVVAISLATFFFLYSFVRDVTHPYA 240
QY 241 RNOQSDFYKIPFIVNKTLPVATLISLVYLAGLAAAYQLYGTKYRRFPWLETWLQ 300
DB 241 RNOQSDFYKIPFIVNKTLPVATLISLVYLAGLAAAYQLYGTKYRRFPWLETWLQ 300
QY 301 CRKQGLSFFPAMVAVYSLCLPMRSEERYFLNMAVQOVHANINENSWNEEVRIMY 360
DB 301 CRKQGLSFFPAMVAVYSLCLPMRSEERYFLNMAVQOVHANINENSWNEEVRIMY 360
QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
DB 361 ISFGIMSLGLSLAVTSIPSVSNALNWRFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
QY 421 EBYRFTYTPNFVLAIVLPISIVIL 444
DB 421 EBYRFTYTPNFVLAIVLPISIVIL 444

RESULT 10
AAU10187 standard; protein; 490 AA.

AAU10187;
16-JAN-2002 (first entry)
Human Six-Transmembrane Protein of Prostate 1, STMPL.
Human; Six-Transmembrane Protein of Prostate 1; STM; 1; prostate cancer;
benign prostatic hyperplasia; acute prostatitis; testicular cancer;
cryptorchidism; testicular disorder; proliferative disorder; lymphoma;
leukaemia; melanoma; ovarian cancer; breast cancer; pancreatic cancer;
liver cancer; lung cancer; cytostatic.

Key Location/Qualifiers
Domain 209..230
Domain 255..273
Domain 304..325
Domain 360..380
Domain 393..415
Domain 432..452
Domain 490..510

XX WO200172962-A2.
XX 04-OCT-2001.
XX 23-MAR-2001; 2001WO-US009410.
XX 24-MAR-2000; 2000US-0191929P.
XX (SAAT/) SAATCIGLU F.
XX Saactciglu F;
XX WPI; 2001-662926/76.
XX N-PSDB; AAS15793, AAS15801, AAS15802.

PT New polynucleotide for the diagnosis, prevention and treatment for
PT prostate and testis disorders, particularly prostate cancer, comprises
PT prostate-specific or testis-specific nucleic acids.

PS Claim 1; Fig 4B; 114bp; English.

CC The invention relates to substantially pure prostate-specific or testis-
CC specific polypeptides and the nucleic acids encoding them. Also included
CC are vectors and host cells expressing the proteins, a transgenic animal
CC expressing the protein, antibodies against the proteins, probes for
CC detecting the nucleic acids, antisense molecules for the nucleic acids
CC and methods of isolating modulators of the proteins. Compounds that
CC modulate the prostate specific or testis specific polypeptide are useful
CC to diagnose, prevent or treat disorders of the testis or prostate
CC particularly prostate cancer, benign prostatic hyperplasia, acute
CC prostatitis, testicular cancer, cryptorchidism, undescended, retractile,
CC ascending or vanished testis. Other proliferative disorders for which the
CC modulators may be used include lymphoma, leukaemia, melanoma, ovarian
CC cancer, breast cancer, pancreatic cancer, liver cancer and lung cancer.
CC The present sequence represents a prostate specific protein, Six-
CC Transmembrane Protein of Prostate 1, STM;1

Sequence 490 AA;

Query Match 97.4%; Score 2290; DB 4; Length 490;
Best Local Similarity 99.8%; Pred. No. 2.3e-238;
Matches 443; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPNGINGIKARKVTGVIGSGDPAKSLTIRLRGCVHVGIS 60
DB 1 MESISMGSPKSLSETCLPNGINGIKARKVTGVIGSGDPAKSLTIRLRGCVHVGIS 60
QY 61 RNPKEASEFPFHVDTTHEDALTKNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
DB 61 RNPKEASEFPFHVDTTHEDALTKNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
QY 121 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQGPXKASRQVYICSNNIQARQVIE 180
DB 121 RINQYPSNAEYLASLPDLSLVKGFNVSAWALQGPXKASRQVYICSNNIQARQVIE 180
QY 181 LARQNFIPIDGLSSAREIENLPLRFTLWRGPVVAISLATFFFLYSFVRDVTHPYA 240
DB 181 LARQNFIPIDGLSSAREIENLPLRFTLWRGPVVAISLATFFFLYSFVRDVTHPYA 240
QY 241 RNOQSDFYKIPFIVNKTLPVATLISLVYLAGLAAAYQLYGTKYRRFPWLETWLQ 300
DB 241 RNOQSDFYKIPFIVNKTLPVATLISLVYLAGLAAAYQLYGTKYRRFPWLETWLQ 300
QY 301 CRKQGLSFFPAMVAVYSLCLPMRSEERYFLNMAVQOVHANINENSWNEEVRIMY 360
DB 301 CRKQGLSFFPAMVAVYSLCLPMRSEERYFLNMAVQOVHANINENSWNEEVRIMY 360
QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
DB 361 ISFGIMSLGLSLAVTSIPSVSNALNWRFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
QY 421 EBYRFTYTPNFVLAIVLPISIVIL 444

Db |||||
421 EYYRFYTPNFVLAIVPSIVLL 444

RESULT 11
AAU10189
ID AAU10189 standard; protein; 419 AA.

XX AAU10189;
AC
XX 16-JAN-2002 (first entry)
DT

XX Human ORF3 of Six-Transmembrane Protein of Prostate 1.
DE

XX Human; Six-Transmembrane Protein of Prostate 1; STM, 1; prostate cancer;
KW benign prostatic hyperplasia; acute prostatitis; testicular cancer;
KW cryptorchidism; testicular disorder; proliferative disorder; lymphoma;
KW leukaemia; melanoma; ovarian cancer; breast cancer; pancreatic cancer;
KW liver cancer; lung cancer; cytostatic; ORF3.

XX Homo sapiens.
OS

XX WO200172962-A2.
PN

XX 04-OCT-2001.
PD

XX 23-MAR-2001; 2001WO-US009410.
PF

XX 24-MAR-2000; 2000US-0191929P.
PR

XX (SAAT/) SAATCIOGLU F.
PA

XX Saatioglu F;
PI

XX WPI; 2001-662926/76.
DR

XX N-PSDB; AAS15813.
DR

XX New polynucleotide for the diagnosis, prevention and treatment for
PT prostate and testis disorders, particularly prostate cancer, comprises
PT prostate-specific or testis-specific nucleic acids.

XX Claim 1; Fig 4K; 114pp; English.
PS

XX The invention relates to substantially pure prostate-specific or testis-
CC specific polypeptides and the nucleic acids encoding them. Also included
CC are vectors and host cells expressing the proteins, a transgenic animal
CC expressing the protein, antibodies against the proteins, probes for
CC detecting the nucleic acids, antisense molecules for the nucleic acids
CC and methods of isolating modulators of the proteins. Compounds that
CC modulate the prostate specific or testis specific polypeptide are useful
CC to diagnose, prevent or treat disorders of the testis or prostate
CC particularly prostate cancer, benign prostatic hyperplasia, acute
CC prostatitis, testicular cancer, cryptorchidism, undescended, retractile,
CC ascending or vanished testis. Other proliferative disorders for which the
CC modulators may be used include lymphoma, leukaemia, melanoma, ovarian
CC cancer, breast cancer, pancreatic cancer, liver cancer and lung cancer.
CC The present sequence is prostate specific protein, Six-Transmembrane
CC Protein of Prostate 1, STM,1, ORF3

XX Sequence 419 AA;
SQ

Query Match 86.6%; Score 2036; DB 4; Length 419;
Best Local Similarity 99.7%; Pred. No. 5.5e-211;
Matches 394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKYVGVIGSGDFAKSLTIRLRGYNHVG 60
DB 1 MESISMGSPKSLSETCLPNGINGIKDARKYVGVIGSGDFAKSLTIRLRGYNHVG 60

QY 61 RNPKFASEFFPHVVDVTHHEDALTKNTIIFVAIHREHYTSLWDLRHLVGRKILLDVSNM 120
DB 61 RNPKFASEFFPHVVDVTHHEDALTKNTIIFVAIHREHYTSLWDLRHLVGRKILLDVSNM 120

QY 121 RINQYPSNAEYIASLFPDSLIVKGFNVSAWALQCPKASROVYICSNIGARQOVIE 180
DB 121 RINQYPSNAEYIASLFPDSLIVKGFNVSAWALQCPKASROVYICSNIGARQOVIE 180
QY 181 LARQNLNIPIDIGSLSSAREIENLPLRLFTLWRGPVVAISLAFEFLLYSFVRDVIHPYA 240
DB 181 LARQNLNIPIDIGSLSSAREIENLPLRLFTLWRGPVVAISLAFEFLLYSFVRDVIHPYA 240
QY 241 RNOQSDPYKIPIEIVNKTLPVIAITLLSLVYLAGLLAAAYQLYGTYKRRFPMLFTWLQ 300
DB 241 RNOQSDPYKIPIEIVNKTLPVIAITLLSLVYLAGLLAAAYQLYGTYKRRFPMLFTWLQ 300
QY 301 CRKQLGLISFFPAMVHVAYSICLPMRSEERYFLNMAYQOVHANIENSWNEEYVRIEMX 360
DB 301 CRKQLGLISFFPAMVHVAYSICLPMRSEERYFLNMAYQOVHANIENSWNEEYVRIEMX 360
QY 361 ISFGIMSLGILSLAVTSPVSNALNWRREFSFIQ 395
DB 361 ISFGIMSLGILSLAVTSPVSNALNWRREFSFIQ 395

RESULT 12
ABG12306

ID ABG12306 standard; protein; 576 AA.

XX ABG12306;
AC

XX 18-FEB-2002 (first entry)
DT

XX Novel human diagnostic protein #12297.
DE

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.
OS

XX WO200175067-A2.
PN

XX 11-OCT-2001.
PD

XX 30-MAR-2001; 2001WO-US008631.
PF

XX 31-MAR-2000; 2000US-00540217.
PR

XX 23-AUG-2000; 2000US-00649167.
PR

XX (HYSE-) HYSEQ INC.
PA

XX Drmanac RT, Liu C, Tang YT;
PI

XX WPI; 2001-639362/73.
DR

XX N-PSDB; AAS76493.
DR

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX Claim 20; SEQ ID NO 42665; 103pp; English.
PS

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations

RESULT 14
AAE02841
ID AAE02841 standard; protein; 450 AA.
XX AC AAE02841;
XX DT 06-AUG-2001 (first entry)
XX DE Human STEAP-2 protein, alternative version.
XX KW Human; cytostatic; antiproliferative; vaccine; gene therapy;
KW six transmembrane epithelial antigen of the prostate-2; STEAP-2;
KW chromosome 7q21; cancer; prostate; colon; bladder; lung; ovarian;
KW pancreatic.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Misc-difference 335..336
FT /note= "Encoded by GACTGAGCT"
FT Misc-difference 375..376
FT /note= "Encoded by AGATGAAGT"
FT Misc-difference 415..416
FT /note= "Encoded by GAGTAAGGC"
FT Misc-difference 445..446
FT /note= "Encoded by ACATGAAGT"
FT Misc-difference 448..449
FT /note= "Encoded by AATTAAATTC"
XX PN WO200140276-A2.
XX PD 07-JUN-2001.
XX PF 06-DEC-2000; 2000WO-US033040.
XX PR 06-DEC-1999; 99US-00455486.
XX PA (UROG-) UROGENESYS INC.
XX PI Afaar DEH, Hubert RS, Raitano AB, Saffran DC, Mitchell SC;
PI Faris M, Jakobovits A;
XX WPI; 2001-367804/38.
DR N-PSDB; AAD07072.
XX PT New STEAP (six transmembrane epithelial antigen of the prostate)
PT proteins, expressed in human cancers, useful for detecting and treating
PT cancer.
XX PS Claim 1; Page 175-176; 187pp; English.
XX CC The present sequence is an alternative version of human six transmembrane
CC epithelial antigen of the prostate (STEAP)-2 protein. STEAP is a member
CC of cell surface serpentine transmembrane antigens. STEAP-2 gene is
CC located on chromosome 7q21 and is used in gene therapy. Inhibiting the
CC development or progression of a cancer (eg. prostate, colon, bladder,
CC lung, ovarian and pancreatic) expressing STEAP or inhibiting growth or
CC killing cells expressing STEAP in a patient, comprises administering a
CC vaccine composition to the patient. Treating a patient with a cancer that
CC expresses STEAP, or inhibiting growth or killing cells expressing STEAP,
CC comprises administering to the patient a vector encoding single chain
CC monoclonal antibody that comprises the variable domains of the heavy and
CC light chains of the monoclonal antibody that specifically binds to STEAP,
CC such that the vector delivers the single chain monoclonal antibody coding
CC sequence to the cancer cells and the encoded single chain monoclonal
CC antibody is expressed intracellularly. Note: This sequence is stated to
CC be the same as that being shown as SEQ ID NO:8 (AAE02781) in figure 9A-9C
CC of the specification. However the present sequence lacks several amino
CC acids at its N-terminal end and has additional amino acids at its C-
CC terminal end when compared with the sequence shown in figure 9A-9C
XX SQ Sequence 450 AA;

Query Match 73.9%; Score 1738; DB 4; Length 450;
Best Local Similarity 100.0%; Pred. No. 1,1e-178;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 119 NMRINQPPESNAEYLASLPDPSLIVKGFNVVSAWALQGPKDASROYIICSNNIQARQOV 178
DB 1 NMRINQPPESNAEYLASLPDPSLIVKGFNVVSAWALQGPKDASROYIICSNNIQARQOV 60
QY 179 IELARQLNFIPIIDIGSLSSAREIENLPRLFTLMRGPVVVAISLATFFLLYSFVRDVIHP 238
DB 61 IELARQLNFIPIIDIGSLSSAREIENLPRLFTLMRGPVVVAISLATFFLLYSFVRDVIHP 120
QY 239 YARNQOSDFYKPIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTRYRRFPPLMETW 298
DB 121 YARNQOSDFYKPIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTRYRRFPPLMETW 180
QY 299 LQCRKQLGLSFFFAWVHVAYSICLPMRRSERYFLINMAYQOVHANIENSWNEEVRWIE 358
DB 181 LQCRKQLGLSFFFAWVHVAYSICLPMRRSERYFLINMAYQOVHANIENSWNEEVRWIE 240
QY 359 MYISFGIMSLGLSLAVTSIPSVSNALNMBEFSFIQSTLGYVALLSTFHVLIYWKRA 418
DB 241 MYISFGIMSLGLSLAVTSIPSVSNALNMBEFSFIQSTLGYVALLSTFHVLIYWKRA 300
QY 419 FEEYYRFTYTPNFEVLALVLPISVILDLQCRYPD 454
DB 301 FEEYYRFTYTPNFEVLALVLPISVILDLQCRYPD 336

RESULT 15
ABP62883
ID ABP62883 standard; protein; 1082 AA.
XX AC ABP62883;
XX DT 14-OCT-2002 (first entry)
XX DE Human polypeptide SEQ ID NO 320.
XX KW Human; vulnerrary; dermatological; neuroprotective; nootropic; cancer;
KW antiparkinsonian; immunostimulant; cytostatic; immunosuppressive;
KW antidiabetic; antiallergic; gene therapy; wound healing; tissue repair;
KW burn; central nervous system disorder; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease; immune disorder;
KW autoimmune disorder; multiple sclerosis; diabetes; allergy.
XX OS Homo sapiens.
XX PN WO200218424-A2.
XX PD 07-MAR-2002.
XX PF 31-AUG-2001; 2001WO-US027093.
XX PR 01-SEP-2000; 2000US-00654935.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Asundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J;
PI Zhao QA, Wang D, Liu C, Drmanac RT, Wehrman T;
XX WPI; 2002-583321/62.
DR N-PSDB; ABQ93362.
XX PT New polynucleotide and polypeptides, useful for treatment and diagnosis
PT of Alzheimer's, Parkinson's, Huntington's, amyotrophic lateral
PT sclerosis, immune deficiencies, cancer, autoimmune disorders, multiple
PT sclerosis, diabetes and allergies.
XX PS Claim 20; SEQ ID NO 320; 284pp + Sequence Listing; English.
XX CC The invention relates to an isolated polynucleotide (I) comprising one of
CC 245 sequences (ABQ93288-ABQ93532). Treating a condition comprising

CC administering to a mammalian subject a composition comprising the protein
CC (II) encoded by (I) (ABP62809-ABP63053) or an antibody (II) to (I).
CC (I), (II) and (III) are useful for diagnostic evaluation of disorders.
CC (I) is useful for gene therapy of diseases and (II) can be used for
CC therapeutic treatment. Diseases that may be treated include wound healing
CC and tissue repair, burns, central nervous system disorders (e.g.
CC Alzheimer's, Parkinson's, Huntington's and amyotrophic lateral
CC sclerosis), immune deficiencies, cancer, autoimmune disorders, multiple
CC sclerosis, diabetes and allergies. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1082 AA;

Query Match 58.0%; Score 1364.5; DB 5; Length 1082;
Best Local Similarity 83.2%; Pred. No. 1.1e-137;
Matches 272; Conservative 0; Mismatches 0; Indels 55; Gaps 1;

QY 120 MRINQYPSNAEYTLASLPDLSLVKGFNVVSAMALQLGPKDASQVYICSNNIQARQCVI 179
Db 1 MRINQYPSNAEYTLASLPDLSLVKGFNVVSAMALQLGPKDASQVYICSNNIQARQCVI 60
QY 180 ELARQLNFIPIDIGSLSSAREIENLPLRLFTLWRGPVVAISLATFFFLYSFVRDVIHPY 239
Db 61 ELARQLNFIPIDIGSLSSAREIENLPLRLFTLWRGPVVAISLATFFFLYSFVRDVIHPY 120
QY 240 ARNQOSDFYKIPLEIVNKTLPPIVAITLISLVLAGLLAAAYQLYYGTYRRPFWLLETWL 299
Db 121 ARNQOSDFYKIPLEIVNKTLPPIVAITLISLVLAGLLAAAYQLYYGTYRRPFWLLETWL 180
QY 300 QCRKQLGLSFFFAVHVAYSLCLPMRRSERVFLNMAVYQOVHANIENTSNNEEVMRIEM 359
Db 181 QCRKQLGLSFFFAVHVAYSLCLPMRRSERVFLNMAVYQOVHANIENTSNNEEVMRIEM 220
QY 360 YISFGIMSLGLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHYLIYGWKRAF 419
Db 221 YISFGIMSLGLSLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHYLIYGWKRAF 245
QY 420 EEEYRFRYTPPNFVALVLPISIVLDL 446
Db 246 EEEYRFRYTPPNFVALVLPISIVLDL 272

Search completed: April 16, 2004, 12:39:11
Job time : 63 secs